# cancer treatment reports

Includes:
Proceedings of the
Symposium on Designs for
Clinical Cancer Research

Volume 64
Number 2-3
February/March 1980
Pages 199-538

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service National Institutes of Health

# cancer treatment reports

**National Cancer Institute,** Vincent T. DeVita, Jr., *Director*

**Division of Cancer Treatment,** Saul A. Schepartz, *Acting Director*

**Editorial Board**

John S. Macdonald, *Editor-in-Chief*
*Associate Editors:*
Martin D. Abeloff
Murray Brennan
Paul A. Bunn
Eli J. Glatstein
Daniel C. Ihde
Marc E. Lippman
Philip A. Pizzo
Charles A. Schiffer
Mary K. Wolpert
Paul V. Woolley
Daniel S. Zaharko

**Editorial Staff**

Jean Griffin Baum, *Managing Editor*
Elaine P. Beck, *Assistant Editor*
Dee Kittner LeRoy, *Assistant Editor*
Virginia T. Ascosi, *Editorial Assistant*
Mary M. Davis, *Editorial Assistant*
Deidre A. Hoehn, *Editorial Assistant*
Florence L. Barnes

**Advisory Board**

Robert S. Benjamin
Isadore Cohn, Jr
Robert L. Goodman
Ken R. Harrap
Edward Henderson
E. Carmack Holmes
Randall K. Johnson
Irwin H. Krakoff
Robert B. Livingston
Charles G. Moertel
John H. Moxley III
William L. McGuire
David A. Pistenma
Sydney E. Salmon
Vaiunaitis K. Vaitkevicius

## Editorial Policy

**The Editorial Board will consider manuscripts dealing with all aspects of preclinical and clinical cancer treatment. These may include such topics as surgery, radiotherapy, chemotherapy, immunotherapy, supportive care, pharmacology, cell biology and kinetics, mechanisms of drug action, and medicinal chemistry.**

**The journal will consist of two parts. One part will consider previously unpublished manuscripts of original work under four main categories:**

1. **Clinical or preclinical research.**
2. **Correspondence regarding manuscripts previously published in *Cancer Treatment Reports*. Letters on other topics will be published at the discretion of the Editorial Board.**
3. **Invited guest editorials and reviews. Unsolicited contributions will be considered on their individual merits.**
4. **Brief communications. Under this category the Editorial Board will consider short contributions which should not exceed five typewritten pages. Four types of communication will be considered in this category.**
   **(a) Ideas relating to cancer treatment.**
   **(b) Preliminary research reports.**
   **(c) Clinical case reports.**
   **(d) Brief clinical trials of novel treatments.**

**All material submitted for consideration in the journal will be subject to review when appropriate by at least two outside reviewers and one member of the Editorial Board. The decision to publish a manuscript is solely the responsibility of the Editorial Board.**

**The second part of this journal will consist of supplements which may include proceedings of meetings and symposia and program information from the Division of Cancer Treatment, National Cancer Institute. Requests to publish information in *Cancer Treatment Reports* supplement will be considered by the Editorial Board.**

## Subscriptions

*Cancer Treatment Reports* is for sale by the Superintendent of Documents, US Government Printing Office, Washington, DC 20402. The yearly subscription rate is $38.35 domestic, with an additional $9.60 for foreign mailing. The prices vary for individual issues. Payment is required in advance; check or money order should be made payable to the Superintendent of Documents. Please specify List ID "CAN" (the identifying code for this journal) when requesting a subscription.

## INSTRUCTIONS TO AUTHORS

**Summary**
Begin the article with a short summary.

**Text**
Submit an original and three copies of the manuscript, typed double spaced. Each page should begin with a new paragraph. Use accepted spelling, abbreviation, and hyphenation found in *Webster's New International Dictionary*. Footnotes should be typed at the bottom of the page on which they appear.

**Tables**
Submit each table on a separate page, typed double spaced. Please use standard abbreviations; otherwise spell out.

**References**
Submit references typed double spaced. Follow the standard form of *Index Medicus:* author, title, name of journal, volume number, page or inclusive pages, and year. In book references give author, title of book, edition (if more than one), city and name of publisher, date of book, volume number, and pages cited. References in text should begin with number 1 and continue consecutively. References to papers presented at meetings should include author; title of paper; title, city, and date of meeting; city and name of publisher; date of publication; volume number; and pages. Papers presented at meetings but not published should be cited in a footnote rather than in the reference section.

**Illustrations**
Photographs and graphs preferably should be clear, glossy prints. Original art work should be done in black ink for good reproduction. Please send 2 copies of each print. Color photographs cannot be used.

**Nomenclature**
Names of compounds should conform to American usage. See *Chemical Abstracts* for proper nomenclature.

**Abbreviations**
Only standard abbreviations are acceptable. If a new or coined abbreviation is used, it must be defined when first mentioned.

**Brief Reports**
Manuscripts published in the Brief Reports section can be no longer than 5 double-spaced pages including references, tables, and a short summary; figures are usually omitted.

**Letters to the Editor**
Letters to the Editor can be no longer than 3 double-spaced text pages and can be submitted with no more than 1 table and 5 references; figures are usually omitted.

**Symposia**
Requests to publish the proceedings of symposia should be addressed to the Managing Editor. A copy of the agenda and a list of speakers should accompany such requests.

**Submission of Manuscripts**
Manuscripts and all inquiries concerning the status of manuscripts should be addressed to: Managing Editor, Cancer Treatment Reports, National Cancer Institute, 7910 Woodmont Ave, Rm 2C16, Landow Bldg, Bethesda, MD 20014. We invite all authors to submit with their manuscripts a short list of possible reviewers. The editorial board will take these names into consideration in selecting those who will be asked to evaluate your manuscript.

**Reprints**
Supplied free of charge: 250 reprints will be sent to the senior author about three weeks after publication of the article.

## ANNOUNCEMENT

The January 1980 issue of *Cancer Treatment Reports* was the last issue in which Mrs. Pamela Jones served as Managing Editor of *Cancer Treatment Reports*. Pam served as Managing Editor of the journal for a full decade, from 1970–1980. She was responsible for the day-to-day management and organization of *Cancer Treatment Reports* and the Editorial Office. She presided over a number of changes in the journal, including the establishment of a charter for the journal, the implementation of symposia issues, and the design of the change in cover format initiated in 1976. She directed the change in journal scope, content, and format that occurred with the evolution of the journal from *Cancer Chemotherapy Reports* to *Cancer Treatment Reports*.

Both the circulation and the quality of the journal have increased dramatically during the 10 years in which Pam served as Managing Editor. Her devotion to turning out a high quality journal resulted from her individual effort that far exceeded the defined responsibilities of Managing Editor. To the Associate Editors who worked with her, she was unfailingly cheerful and helpful and always ready to assist in problems related to the journal. Those of us within the Division of Cancer Treatment with whom she worked thank her for the time and effort that she extended in establishing *Cancer Treatment Reports* as a cancer treatment journal of the highest quality. We wish her well in her new endeavors.

John S. Macdonald, MD

# Contents

Cancer Treatment Reports
Vol. 64, No. 2–3
February/March 1980

Page

ARTICLES

**Phase II Study of Tamoxifen in Patients With Disseminated Malignant Melanoma** 199
*Edward T. Creagan, James N. Ingle, Stephanie J. Green, David L. Ahmann, and Nai-Siang Jiang*
*Mayo Clinic, Rochester, Minnesota*

**Distribution of Adriamycin in Mice Under Conditions of Local Hyperthermia Which Improve Systemic Drug Therapy** 203
*Richard L. Magin, Richard L. Cysyk, and Charles L. Litterst*
*National Cancer Institute, Bethesda, Maryland*

**Inhibition by Selected Anticancer Agents of the Development of Primary Cell-Mediated Immunity Against Allogeneic Tumor Cells in Culture** 211
*Frank Orsini, Cheryl Eppolito, M. Jane Ehrke, and E. Mihich*
*Roswell Park Memorial Institute, Buffalo, New York*

**Double-Blind Comparison of the Antiemetic Effects of Nabilone and Prochlorperazine on Chemotherapy-Induced Emesis** 219
*Nancy Steele, Richard J. Gralla, David W. Braun, Jr, and Charles W. Young*
*Memorial Sloan-Kettering Cancer Center, New York, New York*

**Phase I-II Study of Mitomycin C Topical Therapy for Low-Grade, Low-Stage Transitional Cell Carcinoma of the Bladder: An Interim Report** 225
*M. Dianne DeFuria, R. Bruce Bracken, Douglas E. Johnson, Mark S. Soloway, Claude E. Merrin, Lee Roy Morgan, Harry C. Miller, and Stanley T. Crooke*
*Bristol Laboratories, Syracuse, New York*

**Combination Antifol Therapy: Observations on the Development of Resistant L1210 Leukemic Cells In Vivo** 231
*George P. Browman and Herbert Lazarus*
*Sidney Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts*

**Modified Procarbazine, CCNU, and Vincristine Combination Chemotherapy (PCV 3) in the Treatment of Malignant Brain Tumors** 237
*Victor A. Levin, Michael S. Edwards, Donald C. Wright, Margaret L. Seager, Tana Pischer Schimberg, Jeanette J. Townsend, and Charles B. Wilson*
*University of California, San Francisco*

Page

**Use of Right Atrial Catheter for Prolonged Iv Support in Cancer Patients** 243
*Ram K. Chillar, Mark Farbstein, Owen B. Ellington, M. Ann Turner, John L. Fahey, Daniel U. Riihimaki, Cindy Smith, Georgina H. A. Fleming, Philip K. Zia, and Ernest Beutler*
*City of Hope National Medical Center, Duarte, California*

**Baker's Antifol in Combination With 5-Fluorouracil and Methyl-CCNU in the Treatment of Metastatic Colorectal Cancer: A Southwest Oncology Group Study (Protocol 7764)** 247
*Michael T. Shaw, John D. Bonnet, Henry Wilson, and Lance K. Heilbrun*
*Southwest Oncology Group*

**Clinical Toxic Effects of 2,3-Dihydro-1*H*-imidazo[1,2-*b*]pyrazole (IMPY) With Relevant Pharmacokinetic Parameters** 251
*James A. Neidhart, Alfred E. Staubus, Donn Young, Gerald W. King, and Louis Malspeis*
*The Ohio State University Hospitals, Columbus*

**Neurotoxicity of Adriamycin in Rats: A Low-Dose Effect** 257
*Bernard S. Jortner and Eun-Sook Cho*
*College of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark*

**Skin Rashes Associated With the Administration of the 2-Nitroimidazole, Misonidazole** 263
*M. I. Saunders, S. Dische, H. D. Kogelnik, Rossall Sealy, and I. Lenox-Smith*
*Mt. Vernon Hospital, Middlesex, England*

**Salvage Treatment of Patients Relapsing After Breast Cancer Adjuvant Chemotherapy** 269
*Albert G. Wendt, Stephen E. Jones, and Sydney E. Salmon*
*University of Arizona College of Medicine, Tucson*

**Effect of Hepatic Microsomal Enzyme Inducers on the Metabolism of Misonidazole in Rats** 275
*Dale Shoemaker, Debbie Upton, and John Strong*
*National Cancer Institute, Bethesda, Maryland*

**Vinblastine Given as a Continuous 5-Day Infusion in the Treatment of Refractory Advanced Breast Cancer** 279
*H.-Y. Yap, G. R. Blumenschein, M. J. Keating, G. N. Hortobagyi, C. K. Tashima, and T. L. Loo*
*The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston*

**Phase I Clinical Study of *N*-(Phosphonacetyl)-L-aspartic Acid (PALA)** 285
*Manuel Valdivieso, E. Colleen Moore, Andrew M. Burgess, Jose R. Marti, Joseph Russ, William Plunkett, Ti Li Loo, Gerald P. Bodey, and Emil J Freireich*
*The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston*

Page

**Effect of Adriamycin on the Cell Kinetics of 13762 Rat Mammary Tumors and Implications for Therapy** 293
*Paul G. Braunschweiger and Lewis M. Schiffer*
*Allegheny General Hospital, Pittsburgh, Pennsylvania*

**cis-Dichlorodiammineplatinum(II) With and Without ICRF-159 in Non-Hodgkin's Lymphoma** 301
*Michael P. Corder, Martin Wiesenfeld, Lawrence C. Maguire, J. Thompson Leimert, and Sherrie K. Panther*
*University of Iowa, Iowa City*

**Complications of Peritoneovenous Shunt for Malignant Ascites** 305
*J. Lokich, R. Reinhold, M. Silverman, and J. Tullis*
*New England Deaconess Hospital and Harvard Medical School, Boston, Massachusetts*

**BRIEF REPORTS AND PRELIMINARY COMMUNICATIONS**

**Lack of Clinical Therapeutic Synergism Between Cyclophosphamide, Adriamycin, and cis-Dichlorodiammineplatinum(II) in Metastatic Human Colorectal Carcinoma** 311
*Michael J. O'Connell, Charles G. Moertel, Joseph Rubin. Allan J. Schutt, and Edward T. Creagan*
*Mayo Clinic, Rochester, Minnesota*

**Evaluation of Vitamin E and Selenium Protection Against Chronic Adriamycin Toxicity in Rabbits** 315
*John F. Van Vleet and Victor J. Ferrans*
*Purdue University, West Lafayette, Indiana*

**Complete Remission of Mediastinal Germ-Cell Tumors With cis-Dichlorodiammineplatinum(II) Combination Chemotherapy** 319
*Q. G. C. M. van Hoesel and H. M. Pinedo*
*University of Utrecht, The Netherlands*

**Treatment of Advanced Ovarian Carcinoma With a Combination of Hexamethylmelamine, Cyclophosphamide, Methotrexate, and 5-Fluorouracil (Hexa-CAF) in Patients With and Without Previous Treatment** 323
*J. P. Neijt, A. C. M. van Lindert, C. P. J. Vendrik, K. J. Roozendaal, A. Struyvenberg, and H. M. Pinedo*
*University Hospital, Utrecht, The Netherlands*

**Burkitt's Lymphoma: Tumor Lysis Following Malignant Hyperthermia** 327
*Robert P. Suer, Willian B. Fisher, Richard W. Pearson, and Douglas A. Triplett*
*Ball Memorial Hospital, Muncie, Indiana*

Page

**Vinblastine, Intermittent Bleomycin, and Single-Dose *cis*-Dichlorodiammineplatinum(II) in the Management of Stage III Testicular Cancer** 331
*George J. Bosl, Roger Kwong, Paul H. Lange, Elwin E. Fraley, and B. J. Kennedy*
*University of Minnesota College of Health Sciences, Minneapolis*

**Combination Chemotherapy With VM-26, Adriamycin, Bleomycin, and Prednisone as a Secondary Treatment of Malignant Lymphoma** 335
*A. Goldhirsch, M. Pirovino, R. W. Sonntag, L. Tschopp, H. J. Ryssel, and K. W. Brunner*
*University Hospital, Inselspital Bern, Switzerland*

**Phase II Evaluation of *N*-(Phosphonacetyl)-L-aspartic Acid (PALA) in Metastatic Adenocarcinoma of the Colon or Rectum** 339
*David A. Van Echo, Charles H. Diggs, Mary Scoltock, and Peter H. Wiernik*
*National Cancer Institute, Baltimore, Maryland*

**Phase II Study of Tamoxifen in Patients With Advanced Renal Cell Carcinoma** 343
*John H. Glick, Alan Wein, Susan Torri, Jane Alavi, David Harris, and Harvey Brodovsky*
*University of Pennsylvania, Philadelphia*

**Phase II Evaluation of 4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA) in Patients With Non-Small Cell Lung Cancer** 345
*Ephraim S. Casper, Richard J. Gralla, David P. Kelsen, Ronald B. Natale, Peter Sordillo, and Alan Houghton*
*Memorial Sloan-Kettering Cancer Center, New York, New York*

**Phase II Evaluation of *N*-(Phosphonacetyl)-L-aspartic Acid (PALA) in Patients With Advanced Colorectal Carcinoma** 349
*David S. Carroll, Richard J. Gralla, and Nancy E. Kemeny*
*Memorial Sloan-Kettering Cancer Center, New York, New York*

**LETTERS**

**Cardiomyopathy in Adults After Combination Adriamycin and DTIC** 353
*Gianni Beretta and Fabrizio Villani*
*S. Carlo Hospital, Milan, Italy*

**Report on Drug-Induced Pericarditis** 353
*Misbahuddin Ahmed and Robert E. Slayton*
*Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois*

**Respiratory Dyskinesia Associated With Hexamethylmelamine** 355
*George B. Murray and Donna B. Greenberg*
*Massachusetts General Hospital, Boston*

Page

**Acute Ventricular Fibrillation and Death During Infusion of 4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA)** 356
*Daniel D. Von Hoff, David Elson, Gary Polk, and Charles Coltman, Jr*
*University of Texas Health Science Center at San Antonio*

**Multiple Ventricular Extrasystoles Following Administration of 4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA)** 358
*G. Falkson*
*H. F. Verwoerd Hospital and the University of Pretoria, Republic of South Africa*

**Vincristine-Induced Orthostatic Hypotension: A Prospective Clinical Study** 359
*Nicholas J. DiBella*
*Fitzsimons Army Medical Center, Aurora, Colorado*

**Phenytoin Sodium-Induced Alterations in the Pharmacokinetics of Misonidazole in the Dog** 360
*R. A. S. White and P. Workman*
*University of Cambridge, England*

**Avascular Necrosis of the Femoral Head: Association With Adjuvant Chemotherapy for Breast Carcinoma** 361
*Marjorie Perloff and Gerson J. Lesnick*
*Mt. Sinai School of Medicine, New York, New York*

## Proceedings of the Symposium on Designs for Clinical Cancer Research

**Ethics and Designs: Laetrile Trials as an Example** 363
*Guy R. Newell and Neil M. Ellison*
*National Cancer Institute, Bethesda, Maryland*

**Clinical Considerations in the Design of Clinical Trials** 367
*Stephen K. Carter*
*Stanford University and University of San Francisco, California*

**Use of Prognostic Factors in Analysis of Historical Control Studies** 373
*Edmund A. Gehan, Terry L. Smith, and Aman U. Buzdar*
*The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston*

**Use of Prognostic Factors in Improving the Design and Efficiency of Clinical Trials in Childhood Leukemia: Children's Cancer Study Group Report** 381
*Denis R. Miller, Sanford Leikin, Vincent Albo, Leonard Vitale, Harland Sather, Peter Coccia, Mark Nesbit, Myron Karon, and D. Hammond*
*Children's Cancer Study Group*

**Sequential Clinical Trials in Cancer Research** 393
*Stephen L. George*
*St. Jude Children's Research Hospital, Memphis, Tennessee*

Page

**Application of Restricted Sequential Design in a Clinical Protocol** 399
*Elizabeth I. Thompson*
*St. Jude Children's Research Hospital, Memphis, Tennessee*

**Patient Heterogeneity in Clinical Trials** 405
*Richard Simon*
*National Cancer Institute, Bethesda, Maryland*

**Discussion I: Basic Design Considerations** 411
*Robert Elashoff*

**Operative Procedures: Is Standardization Feasible or Necessary?** 419
*Jerome J. DeCosse, William L. Donegan, Nell Sedransk, Neil F. Jones, and Dann B. Claudon*
*Medical College of Wisconsin, Milwaukee*

**Quality Assurance Programs in Clinical Trials** 425
*Arvin S. Glicksman, Lawrence E. Reinstein, Roger Brotman, and Daniel McShan*
*Rhode Island Hospital, Providence*

**Use of New Biologic Agents in Clinical Immunotherapy** 435
*Brigid G. Leventhal*
*The Johns Hopkins Oncology Center, Baltimore, Maryland*

**Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part I. Introduction** 441
*Arthur H. Schmale*
*University of Rochester Cancer Center, New York*

**Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part II. Assessing Reliability With Equal-Appearing Interval Scales** 445
*Michael L. Feldstein*
*University of Rochester, New York*

**Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part III. Assessing Measurement Techniques in Psychosocial Oncology** 451
*Gary R. Morrow*
*University of Rochester Cancer Center, New York*

**Evaluation of Toxicity: Clinical Issues** 457
*Teresa J. Vietti*
*Washington University School of Medicine and St. Louis Children's Hospital, Missouri*

**Evaluation of Toxicity: Statistical Considerations** 463
*Jay Herson*
*The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston*

Page

**Discussion II: Critical Definitions and Criteria for Evaluation** 469
*David Byar*

**The Research Logic of Radiation Oncology in Combined Modality Therapy** 473
*Philip Rubin and Omar M. Salazar*
*University of Rochester, New York*

**Animal Models of Cancer Immunotherapy: Questions of Relevance** 481
*David W. Weiss*
*Lautenberg Center for General and Tumor Immunobiology, Hebrew University-Hadassah Medical School, Jerusalem, Israel, and Mt. Siani School of Medicine, New York, New York*

**Prognostic Factors in Cancer of the Testis: An Approach to Identification** 487
*Anne Goldman, George Bosl, Kay Johnson, Elwin E. Fraley, and B. J. Kennedy*
*University of Minnesota, Minneapolis*

**Discussion III: Designs to Optimize Combined-Modality Therapy and Selection and Verification of Prognostic Factors** 497
*Norman Breslow*

**Designing for Cancer Clinical Trials: Selection of Prognostic Factors** 499
*Byron Wm. Brown, Jr*
*Stanford University Medical Center, California*

**Designing for Nonparametric Bayesian Survival Analysis Using Historical Controls** 503
*John Van Ryzin*
*Columbia University, New York, New York*

**Comparisons of Cooperative Group Evaluation Criteria for Multiple-Drug Therapy for Breast Cancer** 507
*Hugh L. Davis, Jr, Phyllis Multhauf, and Jerome Klotz*
*University of Wisconsin, Madison*

**Design of Phase II Clinical Trials in Cancer Using Decision Theory** 519
*Richard J. Sylvester and Maurice J. Staquet*
*Institut Jules Bordet, Brussels, Belgium*

**Pilot Studies and New Agent Programs of the Eastern Cooperative Oncology Group** 525
*John M. Bennett and Marvin Zelen*
*Eastern Cooperative Oncology Group*

**Summary Views: A Statistician's Perspective** 533
*D. R. Cox*
*Imperial College, London, England*

**Perspective of the Clinician** 537
*Barth Hoogstraten*
*Southwest Oncology Group*

*The Secretary of Health and Human Services has determined that the publication of this periodical is necessary in the transaction of the public business required by law of this Department. Use of funds for printing this periodical has been approved by the Director of the Office of Management and Budget through January 31, 1983.*

# Phase II Study of Tamoxifen in Patients With Disseminated Malignant Melanoma [1,2]

**Edward T. Creagan,* James N. Ingle, Stephanie J. Green, David L. Ahmann, and Nai-Siang Jiang [3,4]**

## SUMMARY

**Twenty-five evaluable, ambulatory outpatients with disseminated malignant melanoma received tamoxifen at a dose of 40 mg orally each day. None of the patients had an objective regression and approximately 70% had disease progression within 30 days of starting treatment, including two patients with detectable estrogen receptor titers. Median survival time and time to progression were 4 months and 1 month respectively. Performance score was the most significant covariate for survival time ($P < 0.01$) and time to progression ($P = 0.02$).**

[Cancer Treat Rep 64:199–201, 1980]

Endocrine factors may influence the clinical course of malignant melanoma. Females have a consistent survival advantage compared with males; pregnancy has reportedly accelerated the dissemination of disease and objective tumor regressions have occurred following additive and ablative hormonal treatments (1,2). Of particular note are reports showing that 21%–46% of patients with melanoma have low levels of estrogen receptor (ER) activity in tumor tissue (3–6). ER activity strikingly correlates with the frequency of response to tamoxifen (TAM), an antiestrogen, in patients with breast cancer, a tumor with hormonally dependent characteristics (7). In addition, exogenous estrogens significantly increased the size of murine melanoma nodules and enhanced the growth of pulmonary metastases (8). In the same study, an intriguing observation was objective tumor regression following antiestrogen therapy. Therefore, we embarked on a phase II study to assess the efficacy of TAM in patients with advanced malignant melanoma, a disease for which single-agent chemotherapy and polychemotherapy produce meager response rates of approximately 15% (9,10).

## MATERIALS AND METHODS

Twenty-five patients with measurable, advanced malignant melanoma participated in the study. There were 14 males and 11 females, with a median age of 55 years. All candidates were ambulatory outpatients with adequate nutritional status and without significant radiation therapy. Sixteen patients had visceral-dominant disease and nine had soft tissue and/or nodal involvement. Eastern Cooperation Oncology Group (ECOG) performance scores (PSs) were 0–1 (ten patients), 2–3 (14 patients), and 4 (one patient). Ten patients had had no prior chemotherapy. The remaining 15 patients who had received prior chemotherapy with one or more cytotoxic agents had, except for somewhat worse PSs, covariate distributions of age, sex, and dominant metastatic site which were similar to those of the previously untreated patients (table 1).

The specific estrogen-binding activity was determined in tumor cytosol extracts by a charcoal absorption procedure. Tissue samples (stored at –75° C) were thawed, minced, and homogenized at 0–4° C in six volumes (wt/vol) of 0.05 M tris-HCl, 1 mM EDTA, and 10 mM monothioglycerol, pH 8, by use of Polytron PT 10–20 homogenzer (Brinkman). The cytosol was recovered by centrifugation for 1 hour at 150,000 × *g*. Aliquots were

[1]Received Aug 13, 1979; revised Oct 30, 1979; accepted Nov 13, 1979.

[2]Tamoxifen was supplied by the Clinical Research Department of ICI Americas, Inc, Wilmington, DE.

[3]Division of Medial Oncology (*E. T. Creagan, J. N. Ingle,* and *D. L. Ahmann*), Cancer Statistics Unit (*S. J. Green*), and Department of Laboratory Medicine (*N-S. Jiang*), Mayo Clinic, Rochester, MN.

[4]We wish to thank Dr. David O. Toft, Department of Molecular Medicine, Mayo Clinic, for his assistance in this study.

**Reprint requests to:* E. T. Creagan, MD, Division of Medical Oncology, Mayo Clinic, Rochester, MN 55901.

TABLE 1.—Pretreatment characteristics of 25 patients treated with TAM

| | Prior chemotherapy | |
|---|---|---|
| | None | Some |
| No. of evaluable patients | 10 | 15 |
| Median age in yrs | 55.5 | 55 |
| No. of males/females | 6/4 | 8/7 |
| No. of patients with— | | |
| ECOG PS* | | |
| 0–1 | 6 | 4 |
| 2–3 | 4 | 10 |
| 4 | 0 | 1 |
| Dominant metastatic site | | |
| Lung/liver | 4/1 | 6/3 |
| CNS/soft tissue | 1/5 | 1/4 |

*ECOG score: 0 (fully active) to 4 (totally disabled).

removed for protein determinations by the method of Lowry et al (11).

For the binding assay, a series of incubation tubes were prepared in duplicate which contained 0.15 ml of cytosol and 0.1 ml of other ingredients: $3 \times 10^{-6}$ M testosterone, $0.4–4 \times 10^{-9}$ M $^3$H-estradiol (~ 50 curies/nmol, New England Nuclear), and $3 \times 10^{-6}$ M diethylstilbestrol (in background samples only). Testosterone was included to eliminate possible interference by sex steroid-binding globulin (12). The assay tubes were incubated in an ice bath for 16 hours. A charcoal suspension (0.5 ml) which contained 0.25% charcoal (Sigma), 0.025% dextran T70 (Pharmacia), 10 mM tris-HCl, and 1 mM EDTA, pH 8, was then added to each tube. The suspensions were agitated every 5 minutes for 20 minutes and were centrifuged for 5 minutes at $1000 \times g$. The supernatants were transferred to counting vials and mixed with 5 ml of scintillation fluid (42-ml Fisher Scintiprep 1/1000 ml of toluene). Five concentrations of $^3$H-estradiol were used for each assay and total binding sites were determined by Scatchard analysis (13).

Each patient in the study received TAM at a dose of 40 mg/day orally in two divided doses until death or evidence of tumor progression. This dose was selected because of data from 300 patients with advanced breast cancer which showed a higher response rate from 40-mg/day doses than from lower doses (14).

## EVALUATION

Prior to initial treatment and each subsequent course, patients had a complete medical evaluation, chest roentgenogram, and routine hematologic and chemical parameters.

Objective response was defined as a decrease of ≥ 50% in the product of the longest perpendicular diameters of the most clearly measurable indicator lesion. There must have been no new areas of malignant disease or deterioration in weight, symptoms, or PS. Objective improvement included a 25%–49% decrease in the product of the longest perpendicular diameters of the most clearly measurable lesion. Progressive disease was defined as a > 25% increase in the product of any indicator lesion or the appearance of new lesions. Time to disease progression was calculated from the first day of treatment to the date of progression or last followup.

## RESULTS

No patient satisfied the criteria for an objective response. Approximately 70% of the patients had disease progression at the time of their first re-assessment, yielding a median time to progression of 30 days for the entire group. The median survival time was 124 days from the start of treatment for the patients in our study.

Cox covariate analysis of time to progression and survival was performed using the five covariates of age, sex, prior chemotherapy, PS, and visceral versus nonvisceral metastatic sites (15). For survival, the single most influential covariate was PS ($P < 0.01$); ie, patients with a poor PS fared less well than patients with a good PS. Visceral involvement and prior chemotherapy were of marginal significance alone ($P = 0.05$ and 0.05 respectively), but were not significant when adjusted for PS. For time to progression, a good PS was again shown to be an advantage ($P = 0.02$). The only other covariate of interest was age; alone it was not significant, but after adjustment for PS, an advantage for older patients was suggested ($P = 0.07$). Since we observed no objective regressions, the Cox analysis probably reflects the natural history of malignant melanoma.

Tissue for ER was obtained in 14 patients. Only two patients (14%) had detectable titers (3 and 5 fmols/mg of cytosol protein respectively). These values are low and close to the sensitivity limit of the binding assay. The patient with an ER activity level of five is a 31-year-old male with a PS of 0, no prior chemotherapy, and multiple skin lesions. Fulminant progression of melanoma was noted 1 month after starting treatment. Two additional assays for ER were negative on subsequent specimens taken from the same patient. The other patient with a titer of 3 fmols/mg is a 51-year-old male with visceral-dominant disease who progressed at 1 month after starting treatment.

One patient is of interest. A 34-year-old male had malignant melanoma in an inguinal node and biopsy proven regionally recurrent disease. A clearly measurable, firm left-flank mass (6 × 4 cm) subsequently appeared and TAM was started. However, histologically confirmed scrotal lesions then occurred and the patient had disease progression, but continued to take TAM on his own ac-

cord. Three months later the mass was 2 × 3 cm and 4 months thereafter the mass was barely palpable by three different investigators. There was no recurrence of the scrotal lesions. No new areas of disease have subsequently appeared. Tissue for ER was not obtained.

### Toxic Effects

Four of 25 patients complained about mild-moderate nausea and vomiting. However, one patient discontinued TAM because of severe nausea co-incident with the administration of TAM. Another patient complained of some decrease in taste, but this was not clinically significant.

## DISCUSSION

We were unable to demonstrate substantial clinical benefit from TAM at the dose and schedule utilized in this study of malignant melanoma. Our findings were particularly disappointing since some patients were relatively young and had good PSs, no prior chemotherapy, and primarily soft tissue or nodal involvement. However, one of our patients had some apparent tumor regression using a treatment regimen which we considered unorthodox in that it was continued despite unequivocal tumor progression. Obviously, cautious restraint is mandatory in interpreting this observation because of the capricious nature of malignant melanoma.

Much work is ongoing in the area of protein receptors and neoplasia; eg, androgen, progestagen, and glucocorticoid receptors have been described (16). These findings might explain why some patients have responded to such diverse therapies as estrogens, androgens, and progestational agents (1,5,17). Although the ER activity is a theoretically attractive therapeutic discriminant, it may be of little clinical usefulness in the management of patients with malignant melanoma if it cannot be identified in a more substantial number of patients. In fact, if we had employed the criteria for ER positivity as defined by Fisher et al (> 5 fmols/mg) (4), none of our patients would have been "ER-positive." Moreover, two of the 18 patients with melanoma in a previous study who respnded to additive hormonal treatment were ER-negative (5). Nevertheless, the documented objective regressions which have been reported from hormonal treatments in malignant melanoma cannot be discounted and suggest that further investigation in this area is indicated.

## REFERENCES


1. Bodenham DC, and Hale B. Malignant melanoma. *In* Endocrine Therapy in Malignant Disease (Stoll BA, ed). London, England, WB Saunders Co, Ltd, 1972, pp 377–383.
2. Allen EP. Malignant melanoma: spontaneous regression after pregnancy. Br Med J 2:1067, 1955.
3. Posey LE, Morgan LR, Jr, Beazley RM, et al. Letter: Estrogen receptors. JAMA 238:2599, 1977.
4. Fisher RI, Neifeld JP, and Lippman ME. Oestrogen receptors in human malignant melanoma. Lancet 2(7981):337–339, 1976.
5. Fisher RI, Young RC, and Lippman ME. Diethylstilbestrol therapy of surgically non-resectable malignant melanoma. Proc Am Assoc Cancer Res and ASCO 19:339, 1978.
6. Chaudhuri PK, Walker MJ, Beattie CW, et al. Endocrine influences on the growth of human melanoma in vivo. Scientific program. The Society of Surgical Oncology, Inc, Atlanta, GA, April 17–21, 1979, p 65.
7. Kiang DT, and Kennedy BJ. Tamoxifen (antiestrogen) therapy in advanced breast cancer. Ann Intern Med 87:687–690, 1977.
8. Lopez RE, Bhakoo H, Paolini NS, et al. Effect of estrogen on the growth of B-16 melanoma. Surg Forum 29:153–154, 1978.
9. Constanza ME, Nathanson L, Schoenfeld D, et al. Results with methyl-CCNU and DTIC in metastatic melanoma. Cancer 40:1010–1015, 1977.
10. Luce JK, McBride CM, and Frei E, III. Melanoma. *In* Cancer Medicine (Holland JF, and Frei E III, eds). Philadelphia, PA, Lea and Febiger, 1973, p 1839.
11. Lowry OH, Rosebrough NJ, Farr HL, et al. Protein measurement with folin phenol reagent. J Biol Chem 193:265–275, 1951.
12. Ratajczak T, and Hahnel R. Estradiol receptors: influence of plasma proteins on detection and quantitation. J Steroid Biochem 7:741–744, 1976.
13. Scatchard G. The attraction of proteins for small molecules and ions. Ann NY Acad Sci 51:660–672, 1949.
14. Ward HWC, Arthur K, Banks AJ, et al. Anti-oestrogen therapy for breast cancer—a report on 300 patients treated with tamoxifen. Clin Oncol 4:11–17, 1978.
15. Cox DR. Regression models and life tables. J R Stat Soc (B) 34:187–220, 1972.
16. Lippman ME, ed. Proceedings of the John E. Fogarty International Center Conference on Hormones and Cancer, Bethesda, MD, Mar 29–31, 1978. Cancer Res 38:3981–4376, 1978.
17. Ramirez G, Weiss AJ, Rochlin DB, et al. Phase II study of 6α-methyl-pregn-4-ene-3,11,20-trione (NSC-17256). Cancer Chemother Rep 55:265–268, 1971.

# Distribution of Adriamycin in Mice Under Conditions of Local Hyperthermia Which Improve Systemic Drug Therapy[1]

Richard L. Magin,* Richard L. Cysyk, and Charles L. Litterst[2,3]

## SUMMARY

**Local tumor hyperthermia (LTH) (43°C for 60 minutes) enhanced the effectiveness of Adriamycin (ADR) (10 mg/kg iv) administered systemically against the 16C mouse mammary adenocarcinoma. LTH was produced by the local application of 2450 MHz microwaves to sc implanted tumors. During LTH the intratumor temperature uniformity was ± 0.5°C and the rectal temperatures were 36°C–38°C. The simultaneous application of LTH and ADR resulted in tumor growth delays which were greater than those expected from the addition of individual treatments. Enhanced effectiveness of these treatments did not occur if the two treatments were separated by 4 hours. Cures were observed only in the treatment groups that received drug and heat simultaneously. Measurements of the tissue distribution of ADR and its metabolites using $^{14}$C-ADR indicated that the uptake of radioactivity in the heated tumors was not increased by LTH. A greater percentage of the radioactivity in heated tumors was recovered as ADR metabolites and correspondingly less radioactivity was recovered as unchanged ADR.**

[Cancer Treat Rep 64:203–210, 1980]

The application of local hyperthermia to a tumor has been suggested as a method for selectively enhancing the action of chemotherapeutic agents at the tumor site. Recent reviews have summarized the current understanding of the interactions between heat and anticancer drugs (1–3). Promising results obtained by several researchers using the drug Adriamycin (ADR) in combination with 43°C hyperthermia have indicated a greater than additive increase in the cytotoxicity of these two agents both in vitro (4,5) and in vivo (6). Based on in vitro results, increased uptake and accumulation of ADR in heated tissues have been suggested as a mechanism for the observed synergism (4,6). In order to further investigate the interaction between ADR and hyperthermia in vivo, we have conducted studies using the 16/C mammary adenocarcinoma in mice under conditions of moderate (43°C for 60 minutes) local tumor hyperthermia (LTH). The uptake of radiolabeled ADR and its metabolites into heated and control tumors as well as their distribution to other tissues was also measured.

[1]Received Mar 8, 1979; revised Nov 20, 1979; accepted Dec 5, 1979.

[2]Laboratories of Chemical Pharmacology (*R. L. Magin* and *R. L. Cysyk*) and Toxicology (*C. L. Litterst*), Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD.

[3]We wish to thank Donna Powell for her excellent technical assistance.

**Reprint requests to:* Dr. Richard L. Magin, Department of Electrical Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

## MATERIALS AND METHODS

### Animals and Tumors

Female C3H/He mice were obtained from the Developmental Therapeutics Program (DTP), Division of Cancer Treatment (DCT), National Cancer Institute (NCI). Groups of eight mice were housed in plastic cages and were given pelleted feed and water ad libitum. The 16/C mammary adenocarcinoma was obtained through the courtesy of T. H. Corbett of the Southern Research Institute (Birmingham, AL). This highly metastatic tumor arose spontaneously as a mammary adenocarcinoma in C3H/He mice (7). The tumor is only slightly sensitive to moderate LTH, 43°C for 60 minutes (8), but very sensitive to ADR (7). The tumors were transplanted sc on the left flank anterior to the rear leg with an inoculum (0.2 ml) of a tumor brei (1:5 weight to volume in Hanks' balanced salt solution).

## Microwave Hyperthermia Apparatus

A microwave (2450 MHz) system was utilized to generate controlled, localized hyperthermia in the treated tumors. The output from the microwave source (MW 150, ELMED, Inc, Addison, IL) is equally divided into four separate channels. Each channel is regulated by a power controller (Model A, ELTEK Corp, Gaithersburg, MD) which adjusts the microwave power delivered to each tumor so that a constant temperature (± 0.1°C) is maintained. The tumor temperature is sensed by a small diameter (0.13 mm), insulated, thermocouple wire imbedded at the tumor's center. A digital thermometer (BAT8C, Bailey Instruments, Saddle Brook, NJ) displays the tumor's temperature and provides a reference signal for the power regulator. The microwaves are restricted to the volume of the tumor by using a small diameter (3.2 cm) microwave applicator (AT-502/7z, ELMED, Inc) which is directly coupled to the tumor by a 6 g bolus of muscle-equivalent dielectric material. The tumor encapsulation by the bolus reduces the intratumor temperature gradients to ± 0.5°C in tumors < 8 mm in diameter and increases the efficiency of the applicator. A complete description of the design and operation of the microwave heating apparatus has been published (9).

## Drug Dose and Heat Treatment

ADR was obtained from Farmitalia (Milan, Italy) through the DTP, DCT, NCI. ADR (10 mg/kg) was administered iv by tail vein 5–10 minutes after the tumors were heated to 43°C. The tumors were maintained at 43°C for 60 minutes after ADR administration. The mice were treated between Days 7 and 9 following implantation of the tumors at a time when tumors were 30–60 $mm^3$. All mice were anesthetized with chloral hydrate (525 mg/kg ip) (Matheson, Coleman and Bell, Norwood, OH) prior to heat, drug, or sham treatments. The rectal temperatures of the mice were measured before and after each heat treatment to monitor for whole-body hyperthermia. The control and drug-only groups were kept under heat lamps to maintain constant rectal temperatures equivalent to the heated groups. Mice were weighed and tumors were measured twice each week following treatment. Caliper measurements of tumor length (L) and width (W) were used to estimate tumor volume according to the formula:

$$\text{tumor volume} = \frac{\pi}{6} L \times W^2.$$

Tumor growth delay was determined from the control and treatment groups' tumor growth curves on the day when the control tumors' average size first exceeded 1 $cm^3$. The median day of death for each treatment group was recorded and cures were defined as mice surviving 75 days following tumor implantation with no apparent tumor.

## Drug Distribution

Disposition and plasma clearance of ADR were determined using carbonyl$^{14}$C-ADR (2–11 microcuries/mg). The drug was obtained through the courtesy of the DTP, DCT, NCI. Radioactivity and chemical purity were assessed using the thin-layer chromatography (TLC) method outlined below. Most of the radioactivity, 85%–87%, was found to migrate at an $R_f$ value corresponding to that of ADR. Other $R_f$ values contained various amounts of radioactivity that ranged from 0.6% to 2.9% of the applied dose. At various times after drug administration (2, 15, 30, and 60 minutes and 4 and 24 hours), blood samples were collected by axillary bleeding of anesthetized mice and serum was isolated. The mice were killed by cervical dislocation and liver, spleen, kidney, duodenum, heart, lung, and tumor samples were removed and weighed. Samples (100 mg) were digested at 50°C for 2 hours in Protosol (New England Nuclear Corp, Boston, MA). Decoloration, when needed, was accomplished by adding 100 µl of 30% hydrogen peroxide to the sample and heating for 20 minutes at 50°C. Tissue digests were then diluted with a liquid scintillation mixture (Amersham Corp, Arlington Heights, IL) prior to determination of radioactivity in a Packard Model 3390 liquid scintillation spectrometer. Counting efficiency was determined by external standardization and data were corrected for background and counting efficiency.

## Adriamycin Metabolite Profile

ADR metabolite profiles were determined in heated and control tumors by using a group of mice with 16/C mammary tumors implanted in both the left and right flank. The tumor in one flank was used as a nonheated control tumor and the tumor in the other flank was heated as described above. Mice were killed 4 hours after drug administration. Serum, liver, and tumor were analyzed for radioactivity and the ADR metabolite profile was then determined by TLC. Details of this method have been published (10). Briefly, tissue samples were homogenized in cold $CHCl_3$:MeOH (1:1) and the solvent evaporated to dryness under nitrogen. The residue was dissolved in 0.5 ml of cold $CHCl_3$:MeOH and 50–100 µl was applied to a 250-µm silica gel G thin-layer plate. An ADR standard was also spotted on each plate. The plate was developed for approximately 60 minutes in a mixture of $CHCl_3$:MeOH:HAc:$H_2O$ (80:20:14:6). Spots were visualized using long-wave ultraviolet light. Autoradiograms were prepared by exposing the plate to NS-5T X-ray film (Eastman Kodak Co, Rochester, NY) for 19 days prior to developing the film. Spots were quantitated by dividing the plate into portions, each of which was scraped into a liquid scintillation vial. The radioactivity was determined by adding 20 ml of the scintillation mixture as described.

## RESULTS

No early deaths (1–3 days after treatment) occurred due to toxicity of drug, heat, or the combination of these treatments. There were no skin burns or hair loss at the site of local heat treatments and only small changes (< 6%) in animal weights occurred following the treatments. The results of rectal temperature measurements before and after LTH indicated changes of < 1.0°C.

The results of experiments investigating the effect of LTH, ADR, and the combination of the two treatments on the growth of the 16/C mammary adenocarcinoma are summarized in table 1 and figure 1. LTH alone delayed tumor growth by 3.5, 5.0, and 6.0 days compared with controls in three separate experiments, but did not alter the median day of death with respect to controls (table 1). A single dose of ADR (10 mg/kg) resulted in a 10.5-day tumor growth delay in two separate experiments and a doubling of the median day of death (table 1 and fig 1). The combination of LTH and ADR treatments increased the tumor growth delay to 20 and 24.5 days in separate experiments, but did not increase the median day of death over that of the group receiving ADR alone. The effects of these treatments are clearly shown by the tumor growth curves in figure 1. Twenty-five percent of the mice simultaneously treated with LTH and ADR exhibited tumor cures while no cures were seen in any of the control or other treatment groups. The relatively few cures obtained and the result that the median day of death was not increased more than expected from the addition of LTH and ADR treatments are probably due to the existence of metastases at the time of the local hyperthermia treatment or tumor escape from the treatment area. The 16/C mammary adenocarcinoma is known to metastasize to the lungs in > 75% of transplants (7). In experiment 3, two additional groups were included to determine the effect of the sequence of treatment administration when a 4-hour delay between ADR and LTH was allowed. The 4-hour time delay removed the increase in effectiveness observed when the two treatments were given simultaneously. The tumor growth delays for these sequence groups were approximately the same as those observed for the ADR group. There was no significant difference in the average tumor regrowth characteristics between the group receiving drug alone and the groups receiving drug and heat separated by 4 hours. A small improvement in the median day of death compared with the controls was seen in the two sequence groups, but their median day of death was less than that of either the ADR alone groups or ADR and LTH groups.

In an effort to determine if the interaction between LTH and ADR was due to an enhanced uptake and localization of ADR within the heated tumors, a series of experiments using radiolabeled ADR were conducted. Figure 2 shows the plasma clearance curve and tumor uptake of the radiolabeled ADR. These results are similar to those found by Martini et al (11) using another solid tumor model and show no significant difference in accumulation of radioactivity between the heated and nonheated control tumors at times as early as 2 minutes and as long as 24 hours after drug treatment. The plasma clearance curves for both groups are identical. It should be noted that when the disposition of $^{14}$C-ADR was determined in nontumored control mice at 2 minutes, 30 minutes, and 4 hours after treatment, the plasma decay curves and organ distribution were similar to those found in mice carrying the experimental tumor (data not presented). Table 2 shows the tissue distribution data for tumor-bearing animals. No consistently significant difference between control and LTH mice was found in any tissue.

The results of the experiments to determine the metabolite profile of ADR in heated and control tumors and in control liver are shown in table 3. The radioactivity measured in control tumors was primarily due to ADR, although approximately 18% of the radioactivity migrates with an $R_f$ value which would be expected from the more polar adriamycinol. In addition, approximately 5% of the

TABLE 1.—Results of LTH (43° C for 60 mins) and ADR treatments on tumor growth and mouse survival

| Experiment No. | Group | Treatment day | ADR dose (mg/kg) | Tumor growth delay* (days) | Median day of death† (Day) | Cures (total, 8 mice) |
|---|---|---|---|---|---|---|
| 1 | LTH | 9 | — | 5.0 | 31 | 0 |
| 2 | LTH | 7 | — | 6.0 | 31 | 0 |
| | ADR | 7 | 10 | 10.5 | 64 | 0 |
| | LTH + ADR | 7 | 10 | 24.5 | 64 | 1 |
| 3 | LTH | 7 | — | 3.5 | 33 | 0 |
| | ADR | 7 | 10 | 10.5 | 61 | 0 |
| | LTH + ADR | 8 | 10 | 20.0 | 61 | 3 |
| | LTH/4 hrs/ADR | 9 | 10 | 10.5 | 41 | 0 |
| | ADR/4 hrs/LTH | 9 | 10 | 8.0 | 40 | 0 |

*Tumor growth delay was determined on the day when the control tumors' average size 1st exceeded 1 cm$^3$.
†Median day of death for the control groups for all experiments = 32.8 ± 2.5 days.

TABLE 2.—Distribution of radioactivity following administration of $^{14}$C-ADR iv to C3H/He mice bearing the 16/C mammary adenocarcinoma either without (control) or with LTH

| Tissue | Group | Concentration (μg of ADR equivalents/g of tissue [wet weight]) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Time after treatment | | | | | |
| | | 2 mins | 15 mins | 30 mins | 60 mins | 4 hrs | 24 hrs |
| Tumor | Control | 9.4* ± 1.5 | 7.4 ± 1.8 | 7.2 ± 1.5 | 8.5 ± 1.4 | 6.2 ± 1.2 | 4.6 ± 1.1 |
| | LTH | 7.9 ± 0.8 | 9.2 ± 2.8 | 7.6 ± 2.1 | 8.5 ± 1.1 | 6.1 ± 1.6 | 6.1 ± 1.0† |
| Serum | Control | 17.5 ± 1.0 | 4.2 ± 0.5 | 2.8 ± 0.6 | 2.4 ± 0.4 | 1.0 ± 0.1 | 0.2 ± 0.01 |
| | LTH | 15.8 ± 0.9† | 4.4 ± 0.5 | 2.7 ± 0.4 | 2.5 ± 0.3 | 1.0 ± 0.1 | 0.2 ± 0.07 |
| Liver | Control | 63.0 ± 7.5 | 66.4 ± 5.3 | 48.5 ± 6.6 | 42.1 ± 8.1 | 40.5 ± 8.6 | 10.2 ± 0.6 |
| | LTH | 59.7 ± 3.4 | 58.7 ± 6.2 | 44.6 ± 9.9 | 47.5 ± 6.0 | 52.1 ± 6.6† | 11.7 ± 3.2 |
| Kidney | Control | 81.5 ± 9.6 | 63.7 ± 3.9 | 54.2 ± 8.4 | 45.4 ± 5.1 | 29.3 ± 5.0 | 9.8 ± 1.3 |
| | LTH | 89.2 ± 7.0 | 75.7 ± 1.9† | 67.4 ± 1.1† | 63.6 ± 8.9† | 33.6 ± 3.4 | 11.1 ± 2.6 |
| Spleen | Control | 36.7 ± 1.8 | 35.1 ± 6.7 | 22.5 ± 3.6 | 32.5 ± 3.5 | 27.6 ± 7.1 | 22.8 ± 2.1 |
| | LTH | 32.5 ± 9.2 | 32.9 ± 4.8 | 23.9 ± 4.4 | 29.0 ± 3.0 | 31.5 ± 9.7 | 19.8 ± 3.2 |
| Small intestine | Control | 22.7 ± 2.2 | 47.1 ± 14.6 | 76.0 ± 18.5 | 102 ± 34 | 69.1 ± 36.8 | 3.2 ± 0.2 |
| | LTH | 26.7 ± 4.3† | 64.0 ± 17.4 | 45.3 ± 31.7 | 86.7 ± 22.5 | 74.8 ± 21.3 | 3.7 ± 1.4 |
| Heart | Control | 24.1 ± 2.7 | 18.9 ± 1.5 | 18.4 ± 1.8 | 17.1 ± 2.3 | ND‡ | ND |
| | LTH | 25.7 ± 2.0 | 19.1 ± 1.8 | 15.5 ± 3.0 | 15.5 ± 0.9 | ND | ND |
| Lung | Control | 59.7 ± 4.7 | 34.0 ± 3.2 | 28.8 ± 5.6 | 34.8 ± 6.4 | ND | ND |
| | LTH | 63.0 ± 8.1 | 37.0 ± 2.4 | 31.7 ± 14.0 | 27.8 ± 3.0 | ND | ND |

* Data = mean ± SD (total, 5 mice).
† Data for LTH and control animals were significantly different at $P < 0.05$.
‡ ND = not determined.

radioactivity was found at an $R_f$ that would be expected from the less polar aglycone fraction, and a small amount of radioactivity was associated with an unidentified highly polar compound that remains at the origin (3%). Significantly smaller amounts of radioactivity were present as ADR in the heated tumors than were found in the control tumors. The additional radioactivity in the heated tumors was associated with the aglycone fraction and the unknown metabolite, not adriamycinol.

Table 3 also shows data on the distribution of radioactivity among various $R_f$ values from the TLC of liver, a nonheated control tissue. It can be seen that the metabolite profile from the liver is similar to the metabolite profile from the nonheated tumor.

## DISCUSSION

ADR is an anthracycline antibiotic that has been demonstrated to possess a wide range of antitumor activity against animal and human malignancies (12). The results obtained in this study confirm the interaction between heat and ADR reported by previous investigators, although the mechanism of this interaction in these experiments does not appear to be enhanced uptake and accumulation of parent ADR in the heated tumors. Various investigators have reported synergism between hyperthermia and antineoplastic drugs both in vitro (3–5,13–16) and in vivo (2,6,17–19). Some investigators related the synergism between drugs and heat to increased drug concentration or retention in heated tumors, but other investigators reported no such change in drug uptake. Thus, Rochlin et al (17) demonstrated an increased retention of thioTEPA at 38°C versus 25°C in perfused hind limbs of dogs, and Shingleton (18) showed an increased accumulation of mechlorethamine at 42°C versus 31°C in the VX-2 carcinoma implanted in the hind limbs of rabbits. Osieka et al (20), however, found no difference in methyl-CCNU accumulation between control and locally heated (42.5°C) human colon xenografts in mice, and Weinstein et al (21) reported no increase in methotrexate uptake between locally heated (42°C) and control Lewis lung tumors in mice. In vitro studies of heat plus drug treatment show the same interaction on cell proliferation, and the same apparently drug-specific changes in drug localization as are seen with in vivo heat and drug studies (3–5,13–16). Although several investigators have studied the synergistic action of ADR and heat on tumor growth in vivo, only in vitro studies have investigated the cellular localization of this drug as a function of heating (5,13,14). In vitro studies with Chinese hamster cells showed an increased drug-related fluorescence when they were treated with ADR at 43°C versus 37°C, suggesting (4,6) that the synergistic action of heat and ADR is the result of greater drug uptake into heated cells than into control cells. Results of the present investigation show that radioactivity from $^{14}$C-ADR in control and lo-

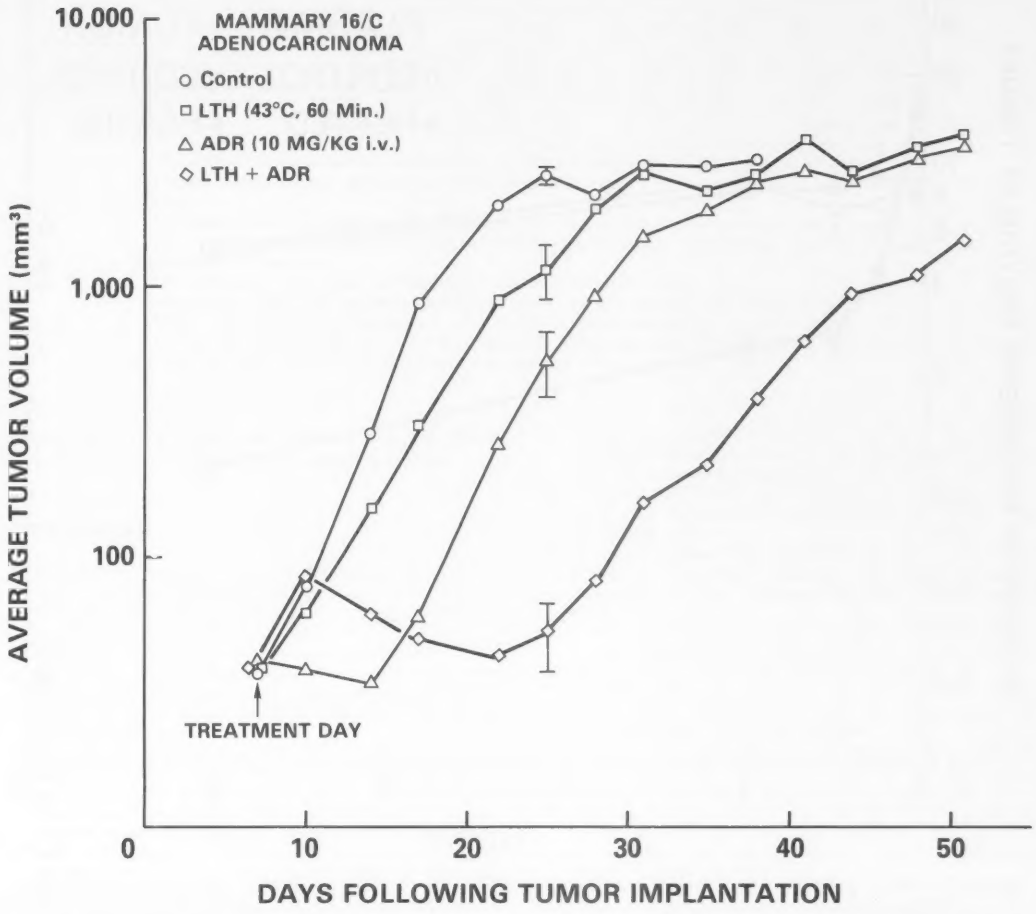



FIGURE 1.—Tumor growth curves for experimental treatment groups as indicated. Each point = average of 8 tumors, and error bars show the SEM. The LTH was administered at 43°C for 60 mins and the ADR was injected by tail vein at a 10-mg/kg dose.

cally heated tumors was equal by 2 minutes after drug administration and remained equal for at least 24 hours. This relatively long retention of the radioactivity within the tumors suggests that the measured tissue radioactivity is due to intracellular ADR or its metabolites bound either to DNA or other cellular components. These findings suggest that the cellular uptake and binding of ADR are not altered by heat treatment. What may be altered in heated tumors is the amount of drug metabolism or the binding of parent drug and metabolites. The data in table 3 show that a greater amount of the intratumor radioactivity in heated tumors is associated with ADR metabolites, particularly the aglycone fraction. Tumors are known to exhibit some drug metabolizing ability, although the existence or extent of this capability in the 16/C mammary adenocarcinoma is unknown. If such enzymes exist, however, they may be expected to be more active at higher temperatures ($Q_{10}$ effect). Alternatively, it may be that binding is altered by hyperthermia and that additional amounts of the smaller aglycone molecule are bound to intratumor binding sites at the expense of the larger parent molecule. Thus, although increased accumulation of the parent drug under conditions of local hyperthermia may account for the reported synergism of heat with some other drugs in vivo, for the tumor and local hyperthermia conditions used in this investigation,

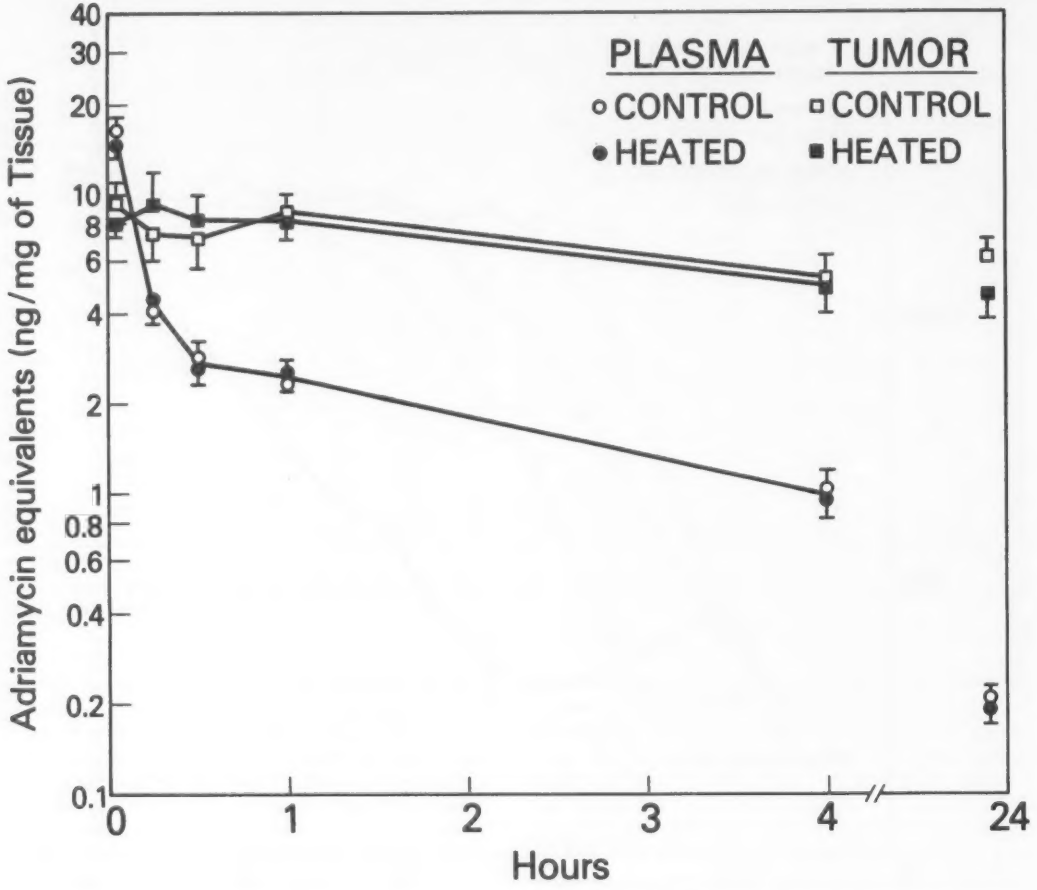



FIGURE 2.—Plasma ADR clearance curves for heated and control mice and tumor ADR uptake curves for heated and control tumors (16/C mammary adenocarcinoma). Each point = average for 5 mice, and the error bars show the SD.

it appears as if a differential distribution and/or retention of ADR metabolites in the heated tumors is responsible for the observed improvement in tumor response.

Bachur et al (22) report higher tissue aglycone concentrations in tissues with active anthracycline metabolism, suggesting the possibility of an increased rate of glycoside bond cleavage in heated tumors. This fact may initially appear inconsistent with efficacy data (table 1) because of the lack of biologic activity of the aglycones. It must be remembered that our TLC system reports total aglycones and thus leaves open the possibility that ADR may be metabolized to adriamycinol and hence to the adriamycinol aglycone at a faster rate in heated tumors. Furthermore, Wilkinson et al (23) have suggested that aglycone formation may be an artifact of the separation method used for ADR and its metabolites which does not occur when a high-performance liquid chromatographic analysis is used, even though aglycone formation has been reported by the same investigators (24) as well as by other investigators (25) using high-performance liquid chromatographic separation techniques. The present results provide no information as to the kind of changes produced in the cell and its receptor molecules by heat treatment. Other investigators have suggested that the anthracyclines are unmetabolized extracellularly and that once metabolized within the cell, the metabolites are retained to a greater degree than the parent drug (26). The data in table 3 are consistent with a heat-related increase in this retention. It appears likely, therefore, that the data shown in table 3 may be the end result of

TABLE 3.—% recovered radioactivity from tissue extracts of locally heated and control tumors migrating at various rates on TLC plates*†

| Compound | $R_f$‡ | Liver | Control tumor | Heated tumor |
|---|---|---|---|---|
| Unknown | 0 – 0.04 | 3.8 ± 1.5§ | 2.7 ± 0.9 | 8.5 ± 3.3¶ |
| Adriamycinol** | 0.04 – 0.28 | 11.7 ± 0.9 | 18.4 ± 0.8 | 16.2 ± 1.1 |
| ADR | 0.28 – 0.43 | 81.2 ± 3.2 | 80.7 ± 3.1 | 71.7 ± 1.7¶ |
| Aglycones†† | 0.43 ± 1.0 | 2.0 ± 1.1 | 5.1 ± 0.8 | 14.5 ± 3.4¶ |

*Total drug concentrations in control (5.88 ± 0.80 μg/g) and heated (5.68 ± 0.87 μg/g) tumors were not significantly different, and control and heated tumor weights were approximately equal for each animal.

†Details of the development and assay systems are presented in the Materials and Methods section.

‡$R_f$ = distance that radioactivity migrated from the origin relative to the distance the solvent front migrated from the origin.

§Data = mean ± SD (total, 4 mice) 4 hrs after ADR administration.

¶Significant differences between control and heated tumors are indicated for $P \leq 0.05$.

**This compound migrated with an $R_f$ value like that expected for the polar metabolite adriamycinol.

††Although an $R_f$ range of 0.43–1.0 is indicated for aglycones, 90%–95% of the radioactivity from all 3 tissues which migrated in this range was localized at $R_f > 0.85$.

complex interactions that take place in the cell during the initial several hours after drug administration.

## REFERENCES


1. HAR-KEDAR I, and BLEEHEN NM. Experimental and clinical aspects of hyperthermia applied to the treatment of cancer with special reference to the role of ultrasonic and microwave heating. Adv Radiat Biol 6:229–266, 1976.
2. HAHN GM. Interactions of drugs and hyperthermia in vitro and in vivo. *In* Proceedings of the 2nd International Symposium on Cancer Therapy by Hyperthermia and Radiation, Essen, Germany, June 2–4, 1977. Baltimore, MD, Urban and Schwarzenberg, 1978, p 344.
3. DICKSON JA, and SUZANGAR M. The in vitro response of human tumors to cytotoxic drugs and hyperthermia (42° C) and its relevance to clinical oncology. *In* Organ Culture in Biomedical Research (Balls M, and Monnickendam MA, eds). London, England, Cambridge University Press, 1976, p 417.
4. HAHN GM, BRAUN J, and HAR-KEDAR I. Thermochemotherapy: synergism between hyperthermia (42–43° C) and Adriamycin (or bleomycin) in mammalian cell inactivation. Proc Natl Acad Sci USA 72:937–940, 1975.
5. HAHN GM, and STRANDE DP. Cytotoxic effects of hyperthermia and Adriamycin on Chinese hamster cells. JNCI 57:1063–1067, 1976.
6. OVERGAARD J. Combined Adriamycin and hyperthermia treatment of a murine mammary carcinoma in vivo. Cancer Res 36:3077–3081, 1976.
7. CORBETT TH, GRISWOLD DP, JR, ROBERTS BJ, ET AL. Biology and therapeutic response of a mouse mammary adenocarcinoma (16/C) and its potential as a model for surgical adjuvant chemotherapy. Cancer Treat Rep 62:1471–1488, 1978.
8. MAGIN RL, and JOHNSON RK. Effects of local tumor hyperthermia on the growth of solid mouse tumors. Cancer Res 39:4534–4539, 1979.
9. MAGIN RL. A microwave system for the controlled production of local tumor hyperthermia in animals. IEEE Trans Microwave Theory Tech 27:78–83, 1979.
10. MIMNAUGH EG, WARING RW, SIKIC BI, ET AL. Effect of whole-body hyperthermia on the disposition and metabolism of Adriamycin in rabbits. Cancer Res 38:1420–1425, 1978.
11. MARTINI A, DONELLI MG, MANTOVANI A, ET AL. Antineoplastic activity and pharmacokinetics of Adriamycin and daunomycin in tumor bearing mice. Oncology 34:173–178, 1977.
12. BLUM RH, and CARTER SK. Adriamycin—a new anticancer drug with significant clinical activity. Ann Intern Med 80:249–259, 1974.
13. BLOCK JB, HARRIS PA, and PEALE A. Preliminary observations on temperature-enhanced drug uptake by leukemic leukocytes in vitro. Cancer Chemother Rep 59:985–988, 1975.
14. KESSEL D, BOTTERILL V, and WODINSKY I. Uptake and retention of daunomycin by mouse leukemic cells as factors in drug response. Cancer Res 28:938–941, 1968.
15. VIG BK. Effect of hypothermia and hyperthermia on the induction of chromosome aberrations by Adriamycin in human leukocytes. Cancer Res 38:550–555, 1978.
16. BRAUN J, and HAHN GM. Enhanced cell killing by bleomycin and 43° C hyperthermia and the inhibition of recovery from potentially lethal damage. Cancer Res 35:2921–2927, 1975.
17. ROCHLIN DB, THAXTER TH, DICKERSON AG, ET AL. The effect of tissue temperature on the binding of alkylating agents in the isolated perfusion treatment of cancer. Surg Gynecol Obstet 113:555–561, 1961.
18. SHINGLETON WW. Selective heating and cooling of tissue in cancer chemotherapy. Ann Surg 156:408–416, 1962.
19. BABA T, AOKI K, YOSHIRO K, ET AL. Temporary interruption of regional blood flow combined with local hyperthermia for cancer chemotherapy. Cancer Res 36:2146–2152, 1976.
20. OSIEKA R, MAGIN RL, and ATKINSON ET. The effect of hyperthermia on human colon cancer xenografts in nude mice. *In* Proceedings of the 2nd International Symposium on Cancer Therapy by Hyperthermia and Radiation, Essen, Germany, June 2–4, 1977. Baltimore, MD, Urban and Schwarzenberg, 1976, p 287.
21. WEINSTEIN JN, MAGIN RL, YATVIN MB, ET AL. Liposomes and local hyperthermia: selective delivery of methotrexate to heated tumors. Science 204:188–191, 1979.
22. BACHUR NR, HILDEBRAND RC, and JAENKE RS. Adriamycin and daunorubicin disposition in the rabbit. J Pharmacol Exp Ther 191:331–340, 1974.
23. WILKINSON PM, ISRAEL M, PEGG WJ, ET AL. Comparative metabolism and excretion of Adriamycin in man, monkey, and rat. Cancer Chemother Pharmacol 2:121–125, 1979.
24. ISRAEL M, WILKINSON PM, PEGG WJ, ET AL. Hepatobiliary metabolism and excretion of Adriamycin and *N*-trifluoroac-

etyl-Adriamycin-14-valerate in the rat. Cancer Res 38:365–370, 1978.

25. LANGONE JJ, VAN VUNAKIS H, and BACHUR NR·Adriamycin and metabolites: separation by high-pressure liquid chromatography and quantitation by radioimmunoassay. Biochem Med 12:283–289, 1975.
26. BACHUR NR. Adriamycin-daunorubicin cellular pharmacodynamics. Biochem Pharmacol (Suppl)2:207–216, 1974.

# Inhibition by Selected Anticancer Agents of the Development of Primary Cell-Mediated Immunity Against Allogeneic Tumor Cells in Culture [1,2]

Frank Orsini,[3,4] Cheryl Eppolito,[3] M. Jane Ehrke,[3] and E. Mihich [3,*]

## SUMMARY

**The effects of 12 anticancer agents on the development of a primary cell-mediated immune response by spleen cells from the C57BL/6J mouse against irradiated allogeneic P815 tumor cells in culture were examined. The agents tested belong to five general classes of compounds: two anthracyclines, Adriamycin and daunorubicin; three vinca alkaloids, vincristine, vinblastine, and vindesine; two structurally related antibiotics, streptozotocin and chlorozotocin; three nitrosoureas, PCNU, BCNU, and methyl-CCNU; and two fluorinated pyrimidines, 5-fluorouracil and ftorafur. These agents, at molar concentrations ranging over at least a tenfold dilution, were added to the primary culture on Day 0 or on Day 2. In general, a given concentration of an agent which inhibited the development of cytotoxic effectors when added on Day 0 of the culture was less effective when added on Day 2. Streptozotocin was ineffective at all concentrations tested when added on either day. Methyl-CCNU and Adriamycin were ineffective when added on Day 2. Within three classes of agents significant differences were seen in the sensitivity of the immune response to the analogs. There was, in fact, a concentration difference of 1 log in the sensitivity of the system to the two fluorinated pyrimidines, 5-fluorouracil being effective at the lower concentration. The development of cytotoxic effectors was also significantly less sensitive to methyl-CCNU than to BCNU at equivalent molar concentrations, and while streptozotocin had no effect, chlorozotocin caused dose-dependent inhibition. For both the vinca alkaloids and the anthracyclines there was no significant difference in effects seen when the individual agents were added on Day 0. When these agents were added on Day 2, however, vinblastine and daunorubicin were effective at significantly lower concentrations than were vindesine and Adriamycin, respectively. Thus, with this sensitization system in culture, it was possible to measure drug effects on the developing immune response in terms of time of addition and concentration; also, this response showed different sensitivity to closely related analogs within a given class of agents. The validity of the information acquired in the culture system has been verified in the mouse.**

[Cancer Treat Rep 64:211–218, 1980]

Despite the increasing efficacy of chemotherapy in the management of certain neoplastic diseases, most of the treatments now in use appear to be limited in their capacity to eradicate the last tumor cells. In general, this is due to the lack of sufficiently selective tumor toxicity of the available anticancer agents.

[1]Received Mar 28, 1979; revised Nov 6, 1979; accepted Nov 13, 1980.

[2]Supported in part by contract N01-CM-57039 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Department of Experimental Therapeutics and Grace Cancer Drug Center, Roswell Park Memorial Institute, New York State Department of Health, Buffalo.

[4]*Present address:* Department of Laboratory Medicine, Roswell Park Memorial Institute, Buffalo, NY.

**Reprint requests to:* Enrico Mihich, MD, Department of Experimental Therapeutics, Roswell Park Memorial Institute, 666 Elm St, Buffalo, NY 14263.

Increased attention is being given to the relationships between chemotherapy and immunity for three basic reasons: (*a*) the existence of indirect evidence suggesting that human tumors have tumor-associated antigens capable of stimulating host defenses (1–3); (*b*) the observation that anticancer drugs cause selective effects on specific components of the immune response in mice (4–6); and (*c*) the potentially exploitable cooperative interactions (as shown in mice) between intrinsically specific but relatively ineffective host immune responses to tumors and the relative nonspecific but highly effective cytoreductive agents available (7). Potential difficulties may derive from the fact that most anticancer agents have been shown to suppress immune responses. A number of recent reports, however, have indicated that the ultimate effect of some of these agents on the host responses need not always be general immunosuppression. In fact, actual augmentation of specific immune responses following exposure to certain anticancer agents has been described (8–11). As information accumulates on the diversified effects anticancer agents may have on the various components of the immune responses, it might become possible to develop chemotherapeutic treatments which circumvent some of the limitations of immunosuppression and actually exploit drug-determined immunomanipulations possibly resulting in advantageous immunoaugmentation.

When various analogs of an active anticancer drug show comparable activity against test tumors, it is difficult to decide which of them should be advanced to clinical tests. In view of the possibility that differences among analogs may exist in terms of their effects on immune responses, and that analogs of choice may be those which exert the least untoward effects on these responses, it is important to develop assay systems which can differentiate between analogs in this respect. To this end, the development in culture of cell-mediated immunity (CMI) of C57BL/6J mice spleen cells against allogeneic P815 mastocytoma was studied as a test system which may be useful in discriminating between the activity of related analogs. The sensitivity of the system was evaluated measuring the effects of 12 anticancer agents, and the results are reported herein. A preliminary account of these studies has been published (12).

## MATERIALS AND METHODS

*Mice.*—C57BL/6J and DBA/2Cr female mice were obtained from the National Cancer Institute's (Bethesda, MD) contractual breeders.

*Tumor.*—P815 mastocytoma was maintained in the ascites form by ip passage of $5 \times 10^6$ cells every 4 days into syngeneic DBA/2Cr mice. The animals were killed by cervical dislocation and the ascitic cells were washed twice in 0.15 M NaCl and once in Roswell Park Memorial Institute (RPMI)-1640 medium and were subsequently used as stimulator cells in culture or in preparation of labeled target cells for the $^{51}Cr$ release test (*see* below). Prior to use for stimulation in culture, the cells, in 3-ml medium on ice, were exposed to 4000 rads of irradiation. The irradiation was delivered with a General Electric Maxitron-300 machine and the conditions were 300 kv, 20 ma, filter of 0.25 mm Cu, source-to-target distance of 50 cm, and an exposure rate of 200 rads/minute.

*Spleen cell suspensions.*—The mice were killed by cervical dislocation and the spleens were excised aseptically and placed in RPMI-1640 medium in sterile Petri dishes on ice. The spleen cells were forced loose with a rubber policeman and washed through a sterile 50-mesh stainless-steel screen. Single cell suspensions were obtained by passage through sterile 200-mesh stainless-steel screens. The cells were washed with medium three times by centrifugation (2000 rpm for 5 minutes at 4°C). The final cell pellet was resuspended in 2 ml of medium and the cells were counted and their viability was assessed by trypan-blue exclusion.

*Stimulation in culture.*—The primary lymphocyte culture technique used (13) was modified in this laboratory as previously reported (10,12). Briefly, $2 \times 10^7$ C57BL/6J spleen cells in 1 ml of culture medium were placed into 35 × 10-mm plastic tissue culture dishes, to which $10^5$ irradiated P815 cells in 0.1 ml were added. The cultures were fed once daily with 0.1 ml of the feeding medium as previously reported (10) and were kept at 37°C in an atmosphere of 7% oxygen, 10% carbon dioxide, and 83% nitrogen for 4 days, until the cells were harvested for assay.

*Agents used.*—The anthracyclines, Adriamycin (ADM) and daunorubicin (DR), were obtained from Dr. L. Lenaz, Adria Laboratories, Inc (Wilmington, DE). The vinca alkaloids, vincristine (VNCR), vindesine (VND), and vinblastine (VNBL); the nitrosoureas, methyl-CCNU, BCNU, and PCNU; the two antibiotics, streptozotocin and chlorozotocin, and the two fluorinated pyrimidines, 5-fluorouracil (5-FU) and ftorafur were supplied by the Division of Cancer Treatment, National Cancer Institute. Dimethyl sulfoxide was used to solubilize chlorozotocin, BCNU, PCNU, and methyl-CCNU and Tween 80 was used to solubilize ftorafur and 5-FU; all the other agents were soluble in RPMI-1640 medium. In every case, however, the final dilutions for testing were made with RPMI-1640 medium, and appropriate controls with solubilizing agents were always included. The agents, at molar concentrations ranging over at least 1 log, were added in 0.1-ml medium to the cultures on Day 0 or on Day 2. In all cases, 0.1 ml of the appropriate control medium was added to the cultures not receiving drugs. In preliminary experiments with each agent, limiting toxicity was determined in terms of number of viable (as determined by trypan-blue exclusion) cells recovered at the end of the 4-day culture, compared to control. All of

the results reported herein were obtained using concentrations of the agents at which the numbers of viable cells recovered were not significantly different from those in the control cultures.

*Cell-mediated immunity assay system: labeling of target cells.*—Washed tumor cells ($5 \times 10^6$) suspended in 0.1-ml medium were incubated in a 50-ml round-bottomed glass centrifuge tube for 20 minutes at 37°C with 50 microcuries of $Na_2\ ^{51}CrO_4$, specific activity, 100–400 millicuries/mg of Cr (Amersham Corp, Arlington Heights, IL). The labeled cells were washed twice with 30-ml cold RPMI-1640 medium, resuspended in 20 ml of medium containing 5% fetal calf serum (FCS), and further incubated for 45 minutes at 37°C. After this incubation, the target cells were pelleted by centrifugation, counted, and resuspended with medium to a final suspension of $2 \times 10^5$ viable cells/ml; 0.1 ml of that was used for each assay tube.

*Effector cells.*—Cells were harvested on the fifth day of culture. A rubber policeman was used to dislodge plastic-adherent cells. The cells were washed with RPMI-1640 medium, counted in the presence of trypan blue, and adjusted to the desired number of viable cells per milliliter with RPMI-1640 medium + 5% FCS.

*Test conditions.*—Varied numbers of effector cells ($2 \times 10^5$—$2 \times 10^6$) in 0.1-ml medium were placed into duplicate 12 × 75-mm plastic tubes and the labeled targets were added (0.1 ml/tube). Final effector to target (E:T) ratios were, therefore, varied from 10:1 to 100:1. After incubation for 4 hours in a humid atmosphere containing 5% $CO_2$, the reaction was stopped by adding 2 ml of ice-cold medium. The cell suspension was centrifuged immediately for 5 minutes at 2000 rpm in a refrigerated (4°C) International PR6000 centrifuge. The supernatant was decanted into 12 × 75-mm disposable glass tubes. The cpm for both the pellet and supernatant were obtained in a gamma spectrometer (Packard model 3375). The results were calculated as:

$$\%^{51}\text{Cr release} = \frac{\text{cpm supernatant}}{\text{cpm pellet} + \text{cpm supernatant}} 100.$$

Percent $^{51}$Cr-specific release represented the percent release using sensitized effector cells minus the percent release using nonsensitized effector cells. Finally, the percent $^{51}$Cr-specific release obtained in the absence of drugs was taken as the control value and set to 100, and all values obtained in the presence of drug were reported in terms of this control. All individual data points within one experiment were expressed as the mean of duplicate assay tubes and all results reported herein are the mean of at least three experiments with standard deviation indicated.

## RESULTS

As shown in figure 1, the three vinca alkaloids studied inhibited the development of the primary response at concentrations ranging from $10^{-8}$ to $10^{-7}$ M. The Day-0 data show an ID50 (concentration causing a 50% inhibition of the response) of $2.2 \times 10^{-8}$ M for VND and an ID50 of $1.9 \times 10^{-8}$ M for both VNCR and VNBL. When these agents were added to the culture 2 days after antigen, less inhibition was produced under these condi-

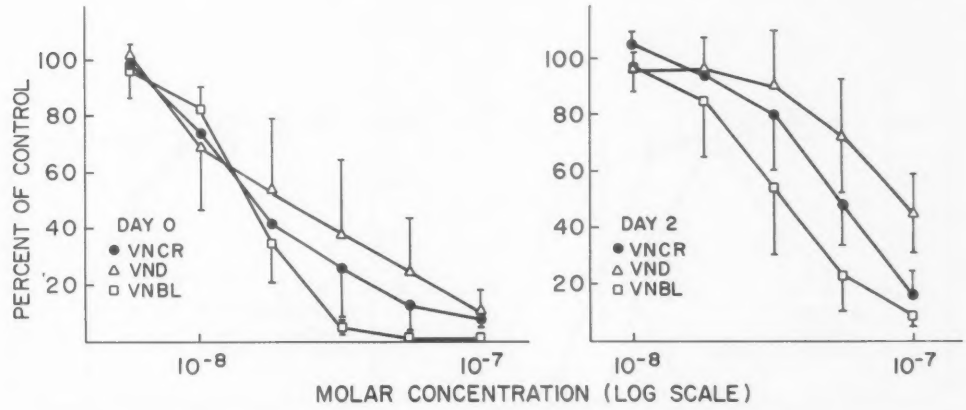



FIGURE 1.—Inhibition of the development of the CMI by VNCR (closed circles), VND (open triangles), and VNBL (open squares) when present in primary sensitization cultures. Drugs were added on either Day 0 (left-hand side) or Day 2 (right-hand side). CMI was measured on the 4th day using a standard 4-hour $^{51}$Cr release assay with an E:T ratio of 100:1. Control = % specific $^{51}$Cr release in the absence of the drug and normalized to 100. Data = a mean of 3 separate experiments, each in duplicate. Vertical lines = SDs.

tions. The ID50s of VND, VNCR, and VNBL were 9 $\times 10^{-8}$ M, 5.5 $\times 10^{-8}$ M, and 3.5 $\times 10^{-8}$ M, respectively. These variations were not significantly different, as indicated by the overlap in standard deviation noted at almost all concentrations studied. VND was different from VNCR and VNBL only when added on Day 2 at a concentration of $10^{-7}$ M.

The two anthracycline antibiotics were suppressive when added on Day 0, with only slight differences in potency (fig 2). The concentrations causing a 50% inhibition of the response were 3 $\times 10^{-7}$ M for ADM and 2.2 $\times 10^{-7}$ M for DR. Both of these agents were nearly ineffective when added on Day 2. Although DR appeared to be more effective than ADM, this difference was significant only at the two highest concentrations tested.

As shown in figure 3, the three nitrosoureas under study, when added on Day 0, were inhibitory only at relatively high concentrations, namely, $10^{-5}$ M to $10^{-4}$ M. The dose-response curve for each drug was very steep, but there were significant differences in the effectiveness of the three compounds, BCNU being the most potent and methyl-CCNU being the least potent. When the agents were added on Day 2, the inhibition of the response was less, and, in fact, methyl-CCNU had no effect

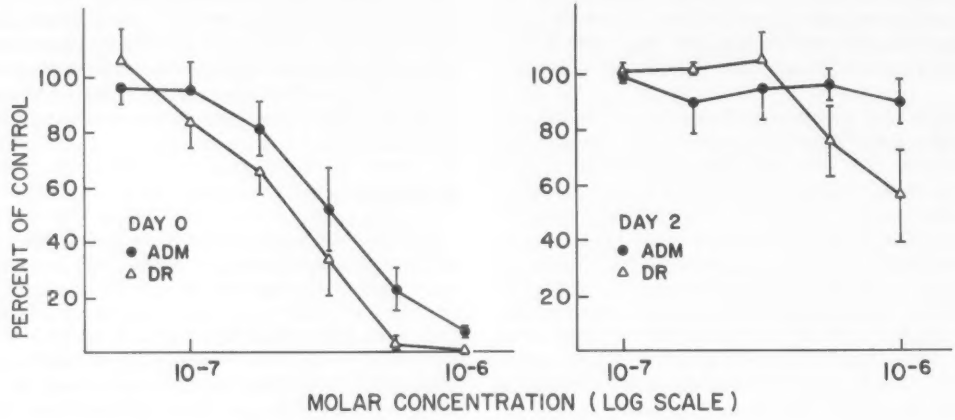



FIGURE 2.—Inhibition of the development of the CMI by ADM (closed circles) and DR (open triangles) when present in primary sensitization cultures. *See* fig 1 for details.

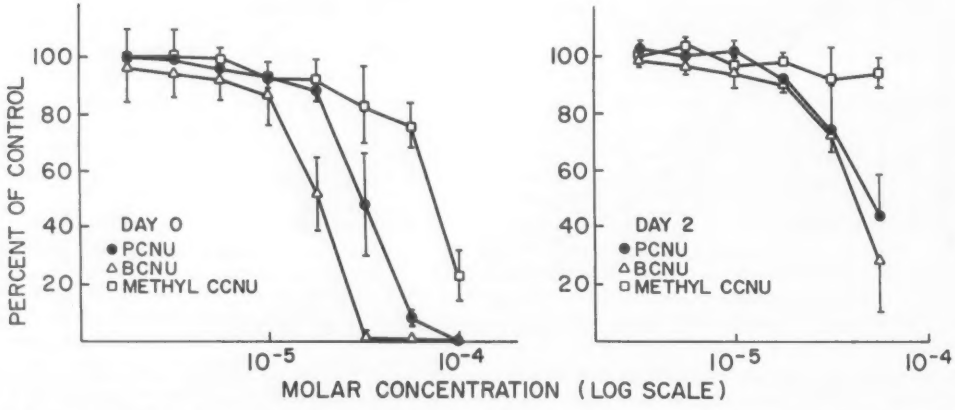



FIGURE 3.—Inhibition of the development of the CMI by PCNU (closed circles), BCNU (open triangles), and methyl-CCNU (open squares) when present in primary sensitization cultures. *See* fig 1 for details.

at noncytotoxic concentrations. The Day-0 ID50s of methyl-CCNU ($7.5 \times 10^{-5}$ M), PCNU ($3.4 \times 10^{-5}$ M), and BCNU ($1.8 \times 10^{-5}$ M) were different.

The effects of streptozotocin and chlorozotocin are shown in figure 4. Streptozotocin had little effect even at $10^{-4}$ M, the highest concentration tested. In contrast, chlorozotocin caused a concentration-dependent inhibition of the response after addition on either Day 0 or Day 2. In this case there was no major difference in effectiveness whether the compound was added on Day 0 (ID50: $2.3 \times 10^{-5}$ M) or on Day 2 (ID50: $5.8 \times 10^{-5}$ M).

A ten to 100 times higher concentration of ftorafur than 5-FU was required to obtain the same inhibition (fig 5). Also, in this case, the inhibition of the response was more marked when the drugs were added to the culture on Day 0 than on Day 2. The ID50 for ftorafur added on Day 0 was $2.5 \times 10^{-5}$ M whereas that of 5-FU was approximately $10^{-6}$ M. The difference between ftorafur and 5-FU was still more marked when added on Day 2 (ID50: $2.3 \times 10^{-4}$ M and $2.8 \times 10^{-6}$ M, respectively).

Table 1 compares the results obtained with these agents in the culture assay system to their effects on the host immune response when given to the mouse. In general, there was a very good agreement between the two

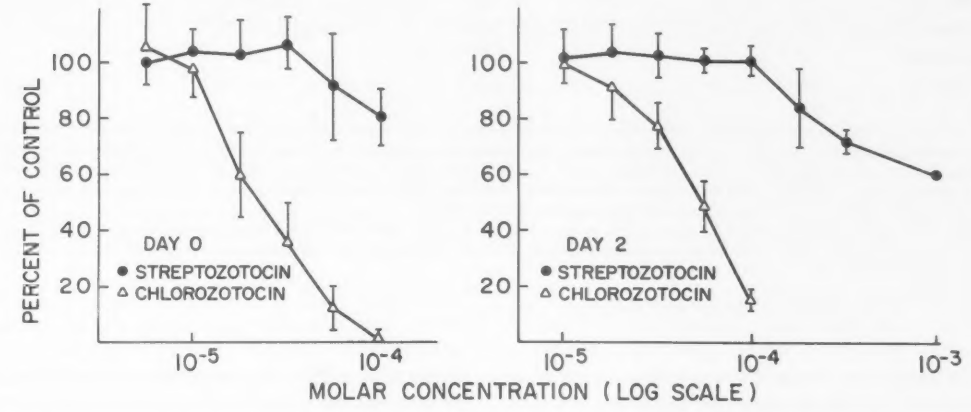



FIGURE 4.—Inhibition of the development of the CMI by streptozotocin (closed circles) and chlorozotocin (open triangles) when present in primary sensitization cultures. *See* fig 1 for details.

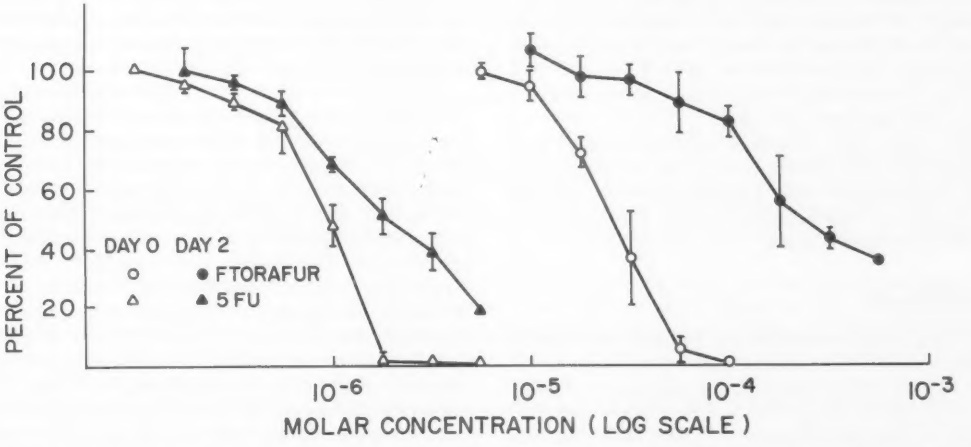



FIGURE 5.—Inhibition of the development of the CMI by ftorafur (open and closed circles) or 5-FU (open and closed triangles) when added on Day 0 (open symbols) or Day 2 (closed symbols) to primary sensitization cultures. *See* fig 1 for details.

TABLE 1.—Comparison of the results obtained with the in vitro test to the effects of the agents in vivo*

| Drug | LD10 (mg/kg) | Drugs given to mice† | | | | Drugs added to culture‡ | | |
|---|---|---|---|---|---|---|---|---|
| | | Day 0 | | Day 2 | | Concentration | | |
| | | LD10 | ½ LD10 | LD10 | ½ LD10 | (M) | Day 0 | Day 2 |
| VNCR | 3.3 | + + + | + | + + | + | $3 \times 10^{-8}$ | + + + | + |
| VND | 5 | NA | NA | NA | NA | $3 \times 10^{-8}$ | + + + | + |
| VNBL | 6.2 | + + + + | + + + + | + + + | + + | $3 \times 10^{-8}$ | + + + + + | + + + |
| ADM | 10 | + + + | – | – | – | $3 \times 10^{-7}$ | + + + | + |
| DR | 10 | + + | – | – | – | $3 \times 10^{-7}$ | + + | – |
| PCNU | 25 | + + + | ND | + + + | ND | $3 \times 10^{-5}$ | + + + | + |
| BCNU | 30 | + + + | ND | + + | ND | $3 \times 10^{-5}$ | + + + + + | + |
| Methyl-CCNU | 33 | + | ND | – | ND | $3 \times 10^{-5}$ | + | – |
| Streptozotocin | 114 | + + | ND | + + + + | + + | $3 \times 10^{-5}$ | – | – |
| Chlorozotocin | 12.5 | + + + + + | + + + + + | + + + + + | + + + + + | $3 \times 10^{-5}$ | + + + | + |
| Ftorafur | 144 | – | – | – | – | $6 \times 10^{-6}$ | – | – |
| 5-FU | 157 | + + + + | + + + | + + + + | + + + + | $6 \times 10^{-6}$ | + + + + + | + + + + |

*An inhibitory effect of the drug (in vivo and in vitro) on the immune response is indicated as follows: + = inhibition of the response of 10%–30%; + + = 30%–50%; + + + = 50%–70%; + + + + = 70%–90%; and + + + + + = > 90%. – = no inhibitory effect with the agent under those conditions when compared to the activity of effector cells from non-drug-treated immunized mice. NA = agent was no longer available. ND = agent was not tested under that condition.

†Drugs were given at the indicated doses on the indicated days. Animals were immunized by ip inoculation of $3 \times 10^7$ P815 cells on Day 0. 12 days later, mice were killed, spleens were excised, single-cell suspension was prepared, and anti-P815 lytic activity was evaluated in the standard 4-hr $^{51}Cr$ release assay.

‡These data are taken from figs 1–5. The concentration at which the results obtained with each drug are shown was selected because it was near or in the ID50 range for that group of agents.

systems when drug treatment was given on Day 0, eg, within each drug group the agent which was the most inhibitory in vitro was also the most inhibitory in vivo (BCNU being the only exception). The comparison of effects resulting from Day-2 drug treatment of the responder mouse or the sensitization culture indicates that treatment at that time was more inhibitory to the response developing in the mouse than to the response developing in culture; however, again there was good agreement between the two systems as to the most active drug within each group of agents.

Table 2 shows that the agents, within the concentration ranges reported in figures 1–5, had no detectable toxic effect on P815 cells, while a number of them, at higher concentrations, could be shown to cause $^{51}Cr$ release from the labeled cells in 4 hours.

## DISCUSSION

Although cancer chemotherapy has achieved a considerable measure of success in that a certain number of patients with certain types of neoplasms can achieve complete remissions through the use of antitumor agents and are free of detectable disease ≥ 5 years after diagnosis, major limitations still need to be overcome before chemotherapy can be broadly used as an effective means to control cancer. These limitations are primarily related to the fact that the drugs available to date, although very potent, are not selective enough in their action against tumor cells. Consequently, even a relatively limited degree of acquired or natural resistance of neoplastic cells to a drug is sufficient to impart clinical resistance because the dose-limiting toxic effects of that drug to normal tissues would forbid the administration of the amounts of drug required to overcome such resistance.

Continuing efforts are being made towards developing new agents with greater antitumor efficacy, namely, greater potency as well as greater selectivity of antitumor action. In many cases, a number of analogs of an active parent structure show similar potency against experimental tumors and the decision must then be made as to which compound to select for advancement towards clinical trials. This selection is customarily made based on a number of parameters, therapeutic index, ie, relative antitumor to lethal effects, being prominent among them. In recent years, however, it has become increasingly evident that preservation of immune responsiveness during treatment with antitumor agents may be critical in determining therapeutic efficacy (7). Indeed, immune responses against tumors may be instrumental in determining the ultimate therapeutic outcome of a treatment and their suppression by chemotherapeutic agents would,

TABLE 2.—Direct toxic effects of the agents on P815 tumor cells in the presence of nonstimulated spleen cells as assessed by a 4-hr $^{51}Cr$ release assay*

| Drug | Concentration (M) ranges tested in— | | Toxicity test | |
|---|---|---|---|---|
| | Experiments shown in figs 1–5 | Toxicity test | Concentration (M) causing $^{51}Cr$ release | % specific $^{51}Cr$ release |
| VNCR | $6 \times 10^{-9}$—$10^{-7}$ | $10^{-7}$—$10^{-4}$ | None | 0 |
| VND | $6 \times 10^{-9}$—$10^{-7}$ | NA† | NA† | NA† |
| VNBL | $6 \times 10^{-9}$—$10^{-7}$ | $10^{-7}$—$10^{-3}$ | $10^{-4}$, $10^{-3}$ | 1, 70 |
| ADM | $6 \times 10^{-8}$—$10^{-6}$ | $10^{-7}$—$10^{-4}$ | $10^{-4}$ | 13 |
| DR | $6 \times 10^{-8}$—$10^{-6}$ | $10^{-7}$—$3.3 \times 10^{-5}$ | $10^{-5}$, $3.3 \times 10^{-5}$ | 1, 45 |
| PCNU | $1.8 \times 10^{-6}$—$10^{-4}$ | $10^{-6}$—$10^{-3}$ | $10^{-3}$ | 2 |
| BCNU | $1.8 \times 10^{-6}$—$10^{-4}$ | $2.5 \times 10^{-8}$—$2.5 \times 10^{-4}$ | $1.5 \times 10^{-4}$, $2.5 \times 10^{-4}$ | 2, 10 |
| Methyl-CCNU | $1.8 \times 10^{-6}$—$10^{-4}$ | ND‡ | ND‡ | ND‡ |
| Streptozotocin | $6 \times 10^{-6}$—$6 \times 10^{-4}$ | $3.3 \times 10^{-7}$—$10^{-2}$ | $10^{-3}$, $10^{-2}$ | 2, 8 |
| Chlorozotocin | $6 \times 10^{-6}$—$10^{-4}$ | $3.3 \times 10^{-7}$—$10^{-2}$ | $10^{-3}$, $10^{-2}$ | 3, 10 |
| Ftorafur | $6 \times 10^{-6}$—$6 \times 10^{-4}$ | $6 \times 10^{-7}$—$10^{-3}$ | None | 0 |
| 5-FU | $10^{-7}$—$6 \times 10^{-6}$ | $6 \times 10^{-7}$—$10^{-4}$ | None | 0 |

*$^{51}Cr$-labeled P815 tumor cells ($2 \times 10^4$) and nonstimulated spleen cells ($10^6$) were mixed with the drugs at the indicated molar concentration ranges in a final volume of 0.2 ml. After 4 hrs incubation the % specific $^{51}Cr$ release was determined as shown in the Materials and Methods section.

†NA = VND no longer available to this laboratory.

‡ND = methyl-CCNU not tested.

therefore, represent an undesirable effect. Moreover, a depression of host defenses by these agents may be responsible, at least in part, for the development of sometimes uncontrollable infections, particularly the so-called opportunistic infections (14). Since suppression of immune responsiveness by drugs need not have a histologically recognizable counterpart, it may escape recognition unless it is measured through functional tests. Thus, the availability of relevant tests of immune function would be particularly useful in the selection of potentially more selective analogs that might have a high ratio for antitumor versus immunosuppressive effect, and consequently might deserve to be tested in man.

The assay system described in this paper was developed to measure the inhibition by anticancer agents of the development in culture of cytotoxic effector cells directed against allogeneic tumor cells. The sensitivity of the lytic effector cell to anti-Thy 1.2 plus complement was established (data not shown). This system may provide information relevant to the potential action of the agents tested on tumor rejection in vivo as T cells appear to play an important role in at least some of the known mechanisms of tumor rejection (15). It must be recognized, however, that when agents are administered in vivo a number of pharmacologic parameters come into play which may have dramatic effects on the overall drug efficacy. This system does not take these parameters into account. Nevertheless, a positive correlation was found between results obtained in culture and those obtained in a comparable primary immunization system in C57BL/6 mice. The main purpose of this study was to test the sensitivity of the assay system in detecting differences in inhibitory potency among closely related anticancer agents as well as among different classes of compounds; 12 known active antitumor agents, belonging to five general classes, were studied to assess the potential usefulness of the system.

When the agents were added on Day 0, significant differences in inhibitory potency were noted among the three nitrosoureas tested (fig 3). Impressive differences were also seen between streptozotocin and chlorozotocin (fig 4), as well as between ftorafur and 5-FU (fig 5). When the drugs were added on Day 2, the anthracyclines (fig 2) and methyl-CCNU (fig 3) were ineffective, while all the other agents were less effective than when added on Day 0. Thus, the assay system studied appears to be sensitive to dose and time of drug addition and capable of detecting differences in potency among related analogs.

While the comparison between the two systems indicated that the Day-2 treatment was inhibitory in the mouse but much less so in culture, it should be noted that the two systems differ in the time required for the peak response to develop (4–5 days in culture [10] and 10–12 days in the mouse [16]); therefore, the Day-2 drug treatment would be relatively late in the response developing in culture but still early in the mouse host response.

All of the drugs tested were effective only during the

early stages of development of the CMI response in culture. Even as little as 12 hours after allogeneic stimulation, ADM and DR were much less effective than when added on Day 0 (17). It is unlikely that the greater inhibition on Day 0 compared to Day 2 is due to drug-caused damage or elimination of the P815 stimulator cells because in the 4-hour $^{51}Cr$ release assay the agents, in the molar concentration ranges tested in the sensitization culture, were shown to cause no increased $^{51}Cr$ release above spontaneous background.

Based on the results of this study it was not possible to identify the cell type or function, the alteration of which might be responsible for the observed effects. The possibility that some of the agents tested (eg, the anthracyclines) act on accessory cells of the monocyte-macrophage type would indeed be consistent with previous results from this laboratory indicating that they selectively affect cells with phagocytic function (5). However, the anthracyclines have been shown to spare macrophages or the potentiality for macrophage repopulation in other systems (10,18). Alternatively, these agents may affect responder cells, responder-cell to stimulator-cell interaction, or the development of precursors of T-cell effectors. The only conclusion that may be reasonably reached in assessing the mechanisms of the inhibition observed is that they are unlikely to be due to a direct antiproliferative effect of the agents studied. In fact, should an indiscriminate antiproliferative action be responsible for the effects reported herein, agents such as ADM or methyl-CCNU would not be so strictly time-dependent in their activity.

In conclusion, the assay system described herein provides useful information on the inhibition of mouse T-cell responses by anticancer agents, but it clearly suffers from the same type of limitations as those of all other preclinical mouse test systems in terms of precise predictions of drug effects in man. The approaches pursued using this assay system, however, may be followed in developing systems based on the responses of human lymphocytes to allogeneic cells or, possibly, to autochthonous tumor cells. If feasible, these immunosuppression assay systems may ultimately be usefully paired with assay systems, possibly predicting drug efficacy against tumor cells such as those recently reported by other investigators (19).

## REFERENCES


1. Herberman RB. Cell mediated immunity to tumor cells. Adv Cancer Res 19:207–264, 1974.
2. Golub SH. Host immune response to human tumor antigens. *In* Cancer, A Comprehensive Treatise (Becker FF, ed). New York, NY, Plenum Press, 1975, vol 4, pp 259–300.
3. Baldwin RW, and Embleton MJ. Assessment of cell-mediated immunity to human tumor-associated antigens. Int Rev Exp Pathol 17:49–95, 1977.
4. Katz SI, Parker D, and Turk JL. B cell suppression of delayed hypersensitivity. Nature 251:550–551, 1974.
5. Ehrke MJ, Cohen SA, and Mihich E. Selectivity of inhibition by anticancer agents of mouse spleen immune effector functions involved in responses to sheep erythrocytes. Cancer Res 38:521–530, 1978.
6. Orsini F, and Mihich E. Immunosuppression by Adriamycin (AM) and daunorubicin (DM). Proc Am Assoc Cancer Res and ASCO 16:130, 1975.
7. Mihich E. Chemotherapy and immunotherapy as a combined modality of cancer treatment. *In* Characterization and Treatment of Human Tumours. (Proceedings of the 7th International Symposium on the Biological Characterization of Human Tumours, Budapest, Hungary, Apr 13–15, 1977). The International Excerpta Medica Congress series No. 420, Advances in Tumour Prevention, Detection, and Characterization. Amsterdam, The Netherlands, Excerpta Medica, 1978, vol 4, pp 113–121.
8. Van Dijk H, and Volermans GAGM. Cyclophosphamide-like cellular immunostimulation induced by 6-thioguanine (Lanvis). Immunology 34:1077–1081, 1977.
9. Turk JL, and Poulter LW. Selective depletion of lymphoid tissue by cyclophosphamide. Clin Exp Immunol 10:285–296, 1972.
10. Orsini F, Pavelic Z, and Mihich E. Increased primary cell-mediated immunity in culture subsequent to Adriamycin or daunorubicin treatment of spleen donor mice. Cancer Res 37:1719–1726, 1977.
11. Orsini FR, and Henderson ES. Increased development of PHA stimulated murine colonies in vitro subsequent to Adriamycin or daunorubicin treatment of spleen donor mice. Immunopharmacology 1:203–210, 1979.
12. Orsini F, Eppolito C, Ehrke MJ, et al. Effects of selected anticancer agents on the development of primary complement-independent cell mediated immune response against allogeneic tumor cells in culture. Fed Proc 37:830, 1978.
13. Cerottini JC, Engers HD, Macdonald HR, et al. Generation of cytotoxic T lymphocytes in vitro. I. Response of normal and immune mouse spleen cells in mixed leukocyte cultures. J Exp Med 140:703–717, 1974.
14. Mihich E. Preclinical evaluation of the interrelationships between cancer chemotherapy and immunity. Natl Cancer Inst Monogr 34:90–116, 1971.
15. Collavo D, Colombatti A, Chieco-Bianchi L, et al. T Lymphocyte requirement for MSV tumor prevention or regression. Nature 249:169–172, 1974.
16. Hoffmann CC, Ehrke MJ, and Mihich E. Temporal specificity of procarbazine in the inhibition of mouse immune responses to allogeneic tumor. J Immunol Pharmacol 1:157–173, 1979.
17. Ehrke J, Tomazic V, Eppolito C, et al. Effects of Adriamycin (AM) in mice or in culture on the primary immune response of C57BL/6 spleen cells to allogeneic P815 mastocytoma cells. Fed Proc 37:1652, 1978.
18. Mantovani A, Tagliabue A, Vecchi A, et al. Effects of Adriamycin and daunomycin on spleen cell populations in normal and tumor allografted mice. Eur J Cancer 12:381–387, 1976.
19. Salmon SE, Hamburger AW, Soehnlen BS, et al. Quantitation of differential sensitivity of human-tumor stem cells to anticancer drugs. N Engl J Med 24:1321–1327, 1978.

# Double-Blind Comparison of the Antiemetic Effects of Nabilone and Prochlorperazine on Chemotherapy-Induced Emesis [1,2,3]

**Nancy Steele, Richard J. Gralla, David W. Braun, Jr, and Charles W. Young [4,*]**

## SUMMARY

**The antiemetic effect of oral nabilone, a synthetic cannabinoid, given at a dose of 2 mg every 12 hours was compared to oral slow-release capsules of prochlorperazine given at a dose of 10 mg every 12 hours by a double-blind crossover method in 37 patients receiving cancer chemotherapy. Patients received one of the following as the primary emetic stimulus: high-dose *cis*-dichlorodiammineplatinum(II) (DDP), low-dose DDP, mechlorethamine, streptozotocin, actinomycin D, or DTIC. Although results varied according to strength of emetic stimulus received, both nabilone and prochlorperazine appeared to produce antiemetic effects. Eighteen of the 37 patients achieved a complete or partial elimination of symptoms: seven with nabilone alone, three with prochlorperazine alone, and eight with each drug. Nabilone appeared to be the more effective antiemetic for patients who received chemotherapy agents *other* than high-dose DDP; it was equivalent to prochlorperazine for those who did receive high-dose DDP. Side effects from prochlorperazine were limited to mild drowsiness occurring among 35% of the patients. The side effects from nabilone were drowsiness and dizziness which occurred frequently and were dose-limiting in 25% of patients.**

[Cancer Treat Rep 64:219–224, 1980]

Nausea and vomiting secondary to cancer chemotherapy are not adequately controlled by antiemetics currently in use. The development of effective and clinically useful antiemetics has become a major goal in the treatment of patients with malignant disease. In 1975, δ-9-tetrahydrocannabinol (THC), the active component of marijuana, was reported to have antiemetic activity against chemotherapy-induced nausea and vomiting (1). Nabilone is a synthetic analog of THC, with a ketone substituted for a methyl group on the C ring and a 1,1-dimethylheptyl side chain (fig 1). Initial pharmacologic studies suggested that nabilone minimized the cardiovascular and psychogenic effects of THC at therapeutic doses (2).

Iv administration of naloxone blocks the effects of nabilone in cats (3), indicating that the drug has a central locus of action which probably involves opiate receptors in the brain. Nabilone does not interfere with apomorphine-induced emesis in the dog.[5] In this, its mode of action appears to differ from that of the phenothiazines which are known to block dopaminergic receptors that are responsible for apomorphine-induced vomiting.

In a phase I trial of nabilone, Herman and co-workers (4) observed that patients refractory to other antiemetics "experienced clinically useful reduction in chemotherapy

[1]Received May 14, 1979; revised Oct 18, 1979; accepted Oct 25, 1979.

[2]Supported in part by Public Health Service grants CA-08748 and CA-05826 and contract N01-CM-57043 (Division of Cancer Treatment) from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Nabilone: compound 109514; *dl*-(1,1-dimethylheptyl)-6,6aβ,7,8,10,10aα-hexahydro-1-hydroxy-6,6-dimethyl-9*H*-dibenzo[*b*,*d*]pyran-9-one. Supplied by Eli Lilly and Co, Indianapolis, IN.

[4]Developmental Chemotherapy Service, Department of Medicine (*N. Steele, R. J. Gralla,* and *C. W. Young*) and Biostatistics Laboratory (*D. Braun, Jr*), Sloan-Kettering Institute for Cancer Research, Memorial Sloan-Kettering Cancer Center, New York, NY.

**Reprint requests to:* Dr. Charles W. Young, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10021.

[5]Stark P. Effects of nabilone (compound 109514) on apomorphine-induced emesis in dogs. Unpublished observations.

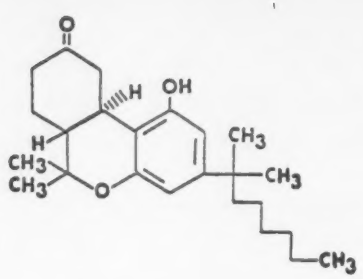


FIGURE 1.—Structural formula of nabilone.

induced nausea and vomiting." In a subsequent crossover study with prochlorperazine, Nagy et al (5) also found that nabilone was an active antiemetic agent. Phase II trials were initiated at multiple centers comparing the efficacy of nabilone and prochlorperazine, a phenothiazine, against nausea and vomiting associated with chemotherapy.

Nabilone trials have been discontinued for the present because of central nervous system toxicity occurring in dogs treated continuously. In spite of this fact, we feel it is useful to offer this report on the data generated in the 9-month trial of nabilone at this Center because of the importance of the clinical problem and the wide interest in cannabinoids as antiemetics.

## MATERIALS AND METHODS

Fifty-five patients at the Memorial Sloan-Kettering Cancer Center were entered in a randomized double-blind crossover comparison study of nabilone and prochlorperazine between April 1978 and January 1979. There were 31 inpatients and 24 outpatients. Patients were not eligible if they had known cardiac disease or psychotic episodes or had regularly used marijuana. Thirty-seven patients ranging in age from 19 to 65 years (median, 50) completed adequate trials of both drugs. Eighteen patients were excluded from evaluation (seven patients had a change in chemotherapy following the first antiemetic drug; three had inadequate data collection; one died prior to the second drug; and four withdrew from the study because of adverse effects and three withdrew because of lack of efficacy).

Subjects received 2 mg of oral nabilone and 10 mg of oral slow-release prochlorperazine in random order with two consecutive and identical chemotherapy treatments. Each antiemetic was given every 12 hours for three to five doses, with the first dose given the night before chemotherapy. Chemotherapy included one of the following as the primary emetic stimulus: high-dose *cis*-dichlorodiammineplatinum(II) (DDP), low-dose DDP, mechlorethamine, streptozotocin, actinomycin D, or DTIC. Table 1 is a complete list of the chemotherapy regimens involved.

A cbc, platelet count, urinalysis, SMA-12, and electro-

TABLE 1.—Combination chemotherapy regimens employed in 37 patients receiving nabilone and prochlorperazine as antiemetic therapy

| Drugs (doses)* | | | | | Frequency |
|---|---|---|---|---|---|
| High-dose DDP (120 mg/m²) | ± | Vindesine (3 mg/m²) | | | q4–6wks |
| Low-dose DDP (60 mg/m²) | ± | Vindesine (3 mg/m²) | | | q4–6wks |
| Low-dose DDP (45–70 mg/m²) | + | Adriamycin (45 mg/m²) | | | q3–4wks |
| Mechlorethamine (6 mg/m²) | + | Vincristine (1.4 mg/m²) | + | Procarbazine orally × 14 days (100 mg/m²) | q4wks (Days 1 and 8) |
| Streptozotocin (500 mg/m²) | | | | | Every wk |
| Actinomycin D (1 mg/m²) | ± | Vinblastine (4 mg/m²) | + | Chlorambucil orally × 14 days (4 mg/m² or 0.15 mg/kg) | q3–4wks |
| DTIC (800 mg/m²) | ± | Cyclophosphamide orally × 14 days (100 mg/m²) | | | q4wks |

*All drugs administere l iv unless otherwise specified.

TABLE 2.—Drug preference of patients

| Preferred antiemetic | No. of patients (%) | | | |
|---|---|---|---|---|
| | High-dose DDP | Low-dose DDP | Other chemotherapy | Total |
| Nabilone* | 8 (53.3) | 6 (60.0) | 9 (75.0) | 23 (62.2) |
| Prochlorperazine (Compazine) | 5 (33.3) | 2 (20.0) | 3 (25.0) | 10 (27.0) |
| Neither | 2 (13.3) | 2 (20.0) | 0 (0.0) | 4 (10.8) |

*% of patients preferring nabilone is not significantly different from 50% in any of the columns ($P > 0.05$ throughout, binomial test).

TABLE 3.—Duration and severity of nausea in relation to the nature of the chemotherapeutic regimen and the antiemetic, nabilone or prochlorperazine

| | High-dose DDP | | Low-dose DDP | | Other chemotherapy | | Total | |
|---|---|---|---|---|---|---|---|---|
| | Nabilone | Prochlor-perazine | Nabilone | Prochlor-perazine | Nabilone | Prochlor-perazine | Nabilone | Prochlor-perazine |
| Duration (days) | | | | | | | | |
| Median | 1.79 | 1.17 | 0.75 | 0.75 | 0.62 | 1.05 | 0.73 | 1.02 |
| Range | 0–3 | 0–2 | 0–2 | 0–2 | 0.5–2 | 0.5–3 | 0–3 | 0–3 |
| Degree of severity* | | | | | | | | |
| Mean | 1.33 | 1.60 | 1.50 | 1.70 | 1.82 | 2.36 | 1.53 | 1.86 |

*Severity scale: 0 = none, 1 = mild, 2 = moderate, and 3 = severe.

cardiogram (ECG) were obtained in hospitalized patients with both test drugs. Supine and standing blood pressures were monitored every 4 hours during waking hours. Outpatients had a cbc, platelet count, and SMA-12 performed before receiving each drug and a baseline urinalysis and ECG were available.

All patients evaluated the therapeutic and adverse effects of both drugs by means of daily questionnaires. Patients rated the severity of nausea (*none; mild:* not interfering with normal activity; *moderate:* interfering with normal activity; and *severe:* bedridden with nausea for 24 hours), estimated the frequency and duration of vomiting, and reported any adverse effects experienced. One investigator (N. Steele) also assessed and recorded each patient's response to therapy based on direct observation and/or interviewing techniques. When both drug studies had been accomplished, patients were asked to state a preference for one drug.

## RESULTS

### Antiemetic Effects

To determine the antiemetic efficacy of each drug, four factors were evaluated and compared: (*a*) patient's drug preference; (*b*) duration and severity of nausea; (*c*) duration and frequency of vomiting; and (*d*) incidence of side effects.

The drug preferences of the patients according to the type of chemotherapy are presented in table 2. In the group of patients who received 120 mg/m$^2$ of DDP by rapid infusion (20-minute), neither antiemetic achieved a significant preference score. Patients who received low-dose DDP by rapid infusion had a 60% preference for nabilone. In each of these groups, 13% and 20% of patients, respectively, did not prefer either drug. All patients in the third group (other chemotherapy) expressed a preference; 75% preferred nabilone and 25% preferred prochlorperazine. Table 2 indicates an overall preference for nabilone, but the difference was not statistically significant ($P > 0.05$, binomial test).

The duration and subjective severity of nausea are recorded in table 3 and the duration and frequency of vomiting are recorded in table 4. The duration of nausea under the influence of nabilone decreased in accord with the expected intensity of emetic stimuli, ie, high-dose DDP > low-dose DDP > other chemotherapy. This was also in accord with the patient's expressed preference. The median duration of nausea did not vary as much with prochlorperazine treatment. The patient's assessments of the severity of nausea did not parallel the changes in duration. All of the patients receiving other types of chemotherapy and one half of the patients receiving low-dose DDP were outpatients who perceived nausea to be moderate to severe more commonly than did the inpatients. The overall duration and severity of nausea was

TABLE 4.—Duration and frequency of vomiting in relation to the nature of the chemotherapeutic regimen and the antiemetic, nabilone or prochlorperazine

| | High-dose DDP | | Low-dose DDP | | Other chemotherapy | | Total | |
|---|---|---|---|---|---|---|---|---|
| | Nabilone | Prochlorperazine | Nabilone | Prochlorperazine | Nabilone | Prochlorperazine | Nabilone | Prochlorperazine |
| Duration (hrs) | | | | | | | | |
| Median | 2.83 | 3.50 | 1.25 | 4.25 | 4.25 | 6.75 | 3.19 | 5.17 |
| Range | 0–48 | 0–36 | 0–7 | 0–7 | 0.5–6 | 0.5–10 | 0–48 | 0–36 |
| Frequency (hrs) | | | | | | | | |
| Median | 7 | 6 | 6.5 | 7.5 | 11 | 18 | 6 | 11.5 |
| Range | 0–35 | 0–continuous | 0–34 | 0–continuous | 2–continuous | 2–continuous | 0–continuous | 1–continuous |

TABLE 5.—Antiemetic drug preference in relation to the order of drug administration

| Order of drug administration | No. of patients preferring— | | Total No. of patients |
|---|---|---|---|
| | Nabilone | Prochlorperazine | |
| First | 9 | 5 | 14 |
| Second | 14 | 5 | 19 |
| Total | 23 | 10 | 33 |

*Only the 33 patients who expressed a drug preference are included. The 19 vs 14 preference for the 2nd drug is not statistically significant (binomial test). There is no statistically significant association between drug preference and the order in which the drugs were received (Fisher's exact test).

slightly less for patients when they were receiving nabilone than when they were receiving prochlorperazine, but the difference was not significant.

The emetic response did not follow the same pattern as the recorded nausea response, that is, decreasing emesis according to the expected decrease in the strength of emetic stimulus. Patients who received "other chemotherapy," all of whom were outpatients, had the highest median duration and frequency of vomiting. For patients receiving high-dose DDP, emesis was not well-controlled with either antiemetic (table 4). Emesis was ablated with only four of 30 treatments: two of 15 courses of nabilone and two of 15 courses of prochlorperazine. The median duration of vomiting for patients receiving low-dose DDP was 3 hours less for patients when they were receiving nabilone than when they were receiving prochlorperazine, but the median frequency was not substantially reduced. Patients who received nonplatinum agents had a decrease in the median duration of vomiting while receiving nabilone; these patients also reported a decreased frequency of emetic episodes with nabilone.

Table 5 presents a cross-tabulation of drug preference for the 33 patients who expressed such by the order in which the drugs were received. There were 14 patients who preferred the first drug they received; for nine patients, that drug was nabilone. There were 19 patients who preferred the second drug they received; for 14 patients that drug was nabilone. The observed preference for the second drug received was not statistically significant (binomial test) and there was no significant association between the drug preferred and the order given when analyzed by the Fisher's exact test.

Antiemetic treatment was instituted the night before chemotherapy and in our trial, 15–18 hours often elapsed before chemotherapy was administered. Because of this pretreatment, a significant number of patients were able to determine which drug they were receiving before chemotherapy was given, by the side effects they experienced. Sixteen patients believed they were receiving nabilone; 13 were correct ($P < 0.05$, binomial test). The three patients who were incorrect preferred this drug believing it was nabilone. This suggests that some degree of blindness was lost; this may have influenced patients' therapeutic response and drug preference in favor of nabilone.

## Side Effects

A total of 53 patients received nabilone at least once (21 patients either did not complete both cycles of the

TABLE 6.—Side effects of antiemetic trials

| | No. of patients | |
|---|---|---|
| | Nabilone (total: 53 patients) | Prochlorperazine (total: 43 patients) |
| Somnolence | 25 (47%) | 15 (34.8%) |
| Dizziness | 19 (35.8%) | 4 (9.3%) |
| Dry mouth | 13 (24.6%) | 2 (4.7%) |
| "High" | 10 (18.9%) | — |
| Postural hypotension | 9 (16.9%) | 2 (4.7%) |
| Increased appetite | 7 (13.2%) | 1 (2.3%) |
| "Drugged" or hangover effect | 5 (9.4%) | — |
| Lightheadedness | 4 (7.5%) | 1 (2.3%) |
| Decreased ability to concentrate | 4 (7.5%) | — |
| Relaxed, tranquil | 3 (5.6%) | — |
| Restlessness | 3 (5.2%) | 2 (4.6%) |
| Nausea | 3 (5.2%) | — |
| Dysphoria | 2 (3.7%) | — |
| Hallucinations | 2 (3.7%) | — |
| Time or space distortion | 2 (3.7%) | — |
| Lethargy | 1 (1.8%) | 1 (2.3%) |
| Headache | 1 (1.8%) | 3 (6.9%) |

study or were otherwise unevaluable). The incidence of side effects for these patients is presented in table 6. Drowsiness, dizziness, and dry mouth occurred in 47%, 36%, and 25% of patients, respectively; they were less annoying to the hospitalized patients. Three outpatients preferred prochlorperazine because the side effects from nabilone outweighed its therapeutic effect. Sixteen outpatients lowered their doses of nabilone to decrease the adverse effects.

Four patients withdrew from the study after one 2-mg dose of nabilone because of intolerable adverse effects. These included severe dizziness, decreased ability to concentrate, lack of coordination, and dysphoria (two patients). One of the dysphoric patients also reported nabilone-induced hallucinations, time and space distortion, and dyspnea. These reactions lasted from 8 to 24 hours and subsided without treatment. Overall, ten patients reported feeling "high" while receiving nabilone. Six of these had had no prior experience with marijuana. Both of the patients experiencing dysphoria were unfamiliar with marijuana.

The incidence of side effects was low for the 43 patients who received prochlorperazine at least once; drowsiness, occurring in 35% of patients, was the only significant effect.

After completing the double-blind study, eight patients continued to receive nabilone on an open-label basis for one to four additional chemotherapy treatments. Their therapeutic and adverse reactions were essentially the same as in the closed-label phase of the study. A higher dose of nabilone was explored in three hospitalized patients. They received one course each of 2 mg of nabilone every 8 hours for six to eight doses. Two patients had the same response as with the previous dose schedule, ie, minimal or no nausea and vomiting with a minimum of side effects. The third patient had a better antiemetic response with this dose and also experienced an increased appetite and a slight euphoric effect.

## DISCUSSION

The present findings provide support for the continuing study of the antiemetic potential of cannabinoids. In the respective drug doses used in this study (2 mg of nabilone and 10 mg of prochlorperazine), nabilone appeared to be the more effective antiemetic for patients who received low-dose DDP and nonplatinum agents. Its effect was equivalent to that of prochlorperazine for those who received high-dose DDP. The type and strength of emetic stimulus received appear to influence the response to antiemetic therapy. Only limited conclusions can be drawn from this comparison because only a single-dose schedule was tested for each drug. The chosen schedule may have been suboptimal for prochlorperazine. The fact that neither nabilone nor prochlorperazine was completely effective against chemotherapy-induced nausea and vomiting supports the need for exploration of other dose schedules.

The side effects from nabilone exceeded those from prochlorperazine and were severe enough to prohibit or modify its use in about 25% of patients. The low incidence of postural hypotension in comparison to previous studies (5,6) may have been due to the number of outpatients in our trial who did not have their blood pressures monitored at home. Eleven outpatients complained of dizziness; one of these patients had a syncopal episode and four were confined to bed because of this side effect.

We did not observe a relationship between nabilone's "euphoric" effect and its antiemetic effect or between the incidence of euphoria and familiarity with marijuana. A subjective "high" was reported by ten patients from 2 mg of nabilone. This reaction was usually mild; it generally started after the first dose of nabilone and lasted for 12–24 hours, regardless of whether therapy was continued or not. Four of these patients received open-label nabilone 2–4 weeks later; two of these patients once again experienced a slight high. One patient took 4 mg of drug and became very high and dizzy for about 18–24 hours.

Patients received a variety of analgesic or sedative medication concurrently with their antiemetic therapy. This combined use of medications could influence the antiemetic results obtained. In this study, the possible bias was lessened by the cross-over design wherein each pa-

tient was entered into both treatment arms. The intensity of analgesic and sedative use was examined and found to be comparable during nabilone and prochlorperazine administration.

Further studies of the antiemetic activity of cannabinoid compounds are indicated. To facilitate comparisons between drugs it may be useful to establish optimal schedules of the current standard antiemetic agents given both by oral and parenteral routes. Since chemotherapeutic emetic stimuli may differ not only in intensity but also in the specific site for triggering emesis, much basic work must be done before comparison of antiemetic agents can be considered routine and reliable.

## REFERENCES


1. Sallan SE, Zinberg NE, and Frei E, III. Antiemetic effect of delta-9-tetrahydrocannabinol in patients receiving cancer chemotherapy. N Engl J Med 293:795–797, 1975.
2. Lemberger L, and Rowe H. Clinical pharmacology of nabilone, a cannabinol derivative. Clin Pharmacol Ther 18:720–726, 1975.
3. McCarthy LE, and Borison HL. Antiemetic activity of nabilone, a cannabinol derivative, reversed by naloxone in awake cats. Pharmacologist 19:239, 1977.
4. Herman TS, Jones SE, Dean J, et al. Nabilone: a potent antiemetic cannabinol with minimal euphoria. Biomedicine 27:331–334, 1977.
5. Nagy C, Furnas B, Einhorn L, et al. Nabilone anti-emetic crossover study in cancer chemotherapy patients. Proc Am Assoc Cancer Res and ASCO 19:30, 1978.

# Phase I–II Study of Mitomycin C Topical Therapy for Low-Grade, Low-Stage Transitional Cell Carcinoma of the Bladder: An Interim Report[1]

M. Dianne DeFuria,[2,*] R. Bruce Bracken,[3] Douglas E. Johnson,[3] Mark S. Soloway,[4] Claude E. Merrin,[5] Lee Roy Morgan,[6] Harry C. Miller,[7] and Stanley T. Crooke[8,9]

## SUMMARY

**Intravesical mitomycin C therapy was administered to 63 patients with noninvasive bladder carcinoma in a phase I–II study. Doses ranged from 20 to 60 mg once a week for 8 weeks. The overall response rate was 67% including a complete remission rate of 45%. The median duration of complete remission for the two lowest dose groups was > 14 months, with a range of 1–36+ months. Followup is not yet adequate to determine the median duration of complete remission for the higher dose groups. No bone marrow suppression or other systemic toxic effects were noted. Adverse effects were localized and occurred infrequently. No mitomycin C was detected in serial serum samples after intravesical instillation.**

[Cancer Treat Rep 64:225–230, 1980]

The polychronotopic nature of many superficial bladder tumors suggests that transurethral resection (TUR) may be insufficient as definitive therapy. Both the field change concept (1) and/or the tumor cell implantation after TUR (2) have led many investigators to employ therapeutic and prophylactic topical chemotherapy of the entire urothelium in addition to surgical management or when surgery is contraindicated.

ThioTEPA has been the agent utilized most frequently in the US and results have been fairly consistent since the early 1960s. Approximately one third of the patients achieve complete disappearance of tumor, one third exhibit a partial response, and one third receive no benefit from therapy (3–6). However, thioTEPA is absorbed from the bladder and significant systemic toxic effects, particularly myelosuppression, have been reported in 12% (7) to 26% (6) of the patients treated. Careful monitoring of bone marrow function with appropriate interruption of therapy to permit recovery is necessary. In addition, a recent report has noted acute myelogenous leukemia occurring in a patient after the intravesical administration of thioTEPA for noninvasive bladder carcinoma (8). Other single agents have been studied with response rates ranging from approximately 17% to 56% (9), but thioTEPA remains the most widely used agent in this country. The value of prophylactic therapy, although suggested to be beneficial, has not been conclusively determined.

In 1975, Mishina et al (10) reported encouraging results, without significant toxicity, using mitomycin C as topical therapy for bladder carcinomas. The study was conducted in 50 patients with stages A–D transitional cell carcinoma. Therapy consisted of a dose of 20 mg of mitomycin C given three times weekly for a total of 20 doses. The response rate was 76% and included a com-

[1]Received May 29, 1979; revised Oct 15, 1979; accepted Oct 25, 1979.

[2]Clinical Cancer Research Department, Bristol Laboratories, Syracuse, NY.

[3]Department of Urology, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston.

[4]Department of Urology, The University of Tennessee College of Medicine, Memphis.

[5]Department of Urologic Oncology, Roswell Park Memorial Institute, Buffalo, NY, and Department of Urology, Loyola University of Chicago, IL.

[6]Department of Pharmacology, Louisiana State University Medical Center, New Orleans.

[7]Department of Urology, The George Washington University Medical Center, Washington, DC.

[8]Research and Development, Bristol Laboratories, Syracuse, NY, and Department of Pharmacology, Baylor College of Medicine, Houston, TX.

[9]We wish to thank Drs. Archie Prestayko, Steven Reich, and Brian Issell for their review of the manuscript and Sue Ellen Briggs for typographical assistance.

*_Reprint requests to:_ M. Dianne DeFuria, MD, PhD, Clinical Cancer Research Department, Bristol Laboratories, Syracuse, NY 13201.

plete remission rate of 44%. The best responses were noted in patients with stage A papillary carcinomas of < 1 cm, but responses occurred regardless of the number of tumors present. Adverse reactions occurred infrequently and were localized. There was no evidence of bone marrow suppression or other systemic toxicity. These data supported previous Japanese studies (11–14) demonstrating the activity of topical mitomycin C and its minimal absorption from the bladder (15).

The present study was initiated to determine the activity and incidence of adverse effects of mitomycin C at various doses administered weekly for 8 weeks in low-grade, low-stage bladder cancer. The extent of drug absorption from the bladder was also determined.

## PATIENTS AND METHODS

Sixty-three informed and consenting adults with biopsy-proven low-grade (I–II) transitional cell carcinoma of the urinary bladder limited to the lamina propria (Marshall stages 0 and A/International Union Against Cancer stages $T_{IS}$ and $T_1$) were entered at five institutions. The majority of patients were treated at the M.D. Anderson Hospital and Tumor Institute (16). There were 48 males and 15 females ranging in age from 32 to 97 years, with a median age of 66 years.

Most patients had undergone multiple surgical procedures with grossly visible residual tumor present at enrollment. Approximately 30% of the patients had received prior chemotherapy, primarily with intravesical thioTEPA. In addition, two patients had had prior radiation therapy. Seventy-three percent of the patients entered in the study had < 25% of the bladder surface involved with tumor. There were also 17 patients with moderate or extensive disease (table 1).

Patients were excluded from the study if the neoplasm extended beyond the lamina propria. Patients with inadequate bone marrow reserve (platelet count < 150,000/mm$^3$ or wbc count < 3500/mm$^3$) or a serum creatinine level of > 2.0 mg/100 ml were also excluded from the study.

TABLE 1.—Patient data

| Characteristic | No. of patients (%) |
|---|---|
| Sex | |
| Males | 48 (76) |
| Females | 15 (24) |
| Age in yrs | |
| Median: 66 | |
| Range: 32–97 | |
| Extent of disease (% surface involvement) | |
| Limited (< 25%) | 46 (73) |
| Moderate (25%–50%) | 13 (21) |
| Extensive (≥ 50%) | 4 (6) |
| Prior therapy | |
| Surgery | 61 (96) |
| Chemotherapy | 18 (29) |
| Radiation | 2 (3) |

FIGURE 1.—Bladder map.

A complete medical history, a physical with bimanual examination, and an endoscopic assessment of disease were performed prior to therapy. Excretory urogram, urinalysis, urine culture and sensitivities, chest roentgenogram, liver profile, and renal function studies were also performed. Bladder maps (fig 1) and tumor measurements, when possible, were made by each investigator. To ensure uniformity, the extent of disease by percent of the mucosal surface involved in tumors was assessed from the diagrams by one individual.

Cbcs and platelet counts were performed prior to and periodically throughout induction therapy. Blood samples were obtained at 0, 60, 120, and 150 minutes after instillation, ie, before and during instillation, and 30 minutes after micturition. The serum samples and a voided urine specimen were frozen for subsequent microbiologic assay. The standard cup agar diffusion technique using either a *Bacillus subtilis* ATCC 6633 spore suspension (17) or a Fujita (18) strain of *Escherichia coli* B was performed. The limit of assay sensitivity was organism-dependent: *B subtilis*, 100 ng of mitomycin C/ml; and *E. coli* B, 10 ng of mitomycin C/ml.

Patients were instructed to abstain from liquids for 12 hours prior to therapy. The 20–40-mg doses of mitomycin C were dissolved in sterile water to a final concentration of 1 mg/ml. The pH was 7.2 and the osmolarity was 10 mOsm. The 60-mg dose had a final concentration of 2.0 mg/ml, a pH of 8.0, and an osmolarity of 25 mOsm. A urethral catheter was inserted into the bladder, the urine was drained completely, and the mitomycin C was instilled. Patients were asked to retain the drug intravesically for 2 hours. The total volume of urine voided was

measured and recorded. This instillation procedure was repeated weekly for 8 weeks. Local adverse reactions were assessed by cystoscopy just prior to the last instillation.

Response to therapy was evaluated by cystoscopy and examination under anesthesia 4 weeks after the last instillation. Biopsies were performed at the discretion of the investigator. Complete remission consisted of the disappearance of all grossly evident disease. Partial remission was considered to be a > 50% decrease in the total area of bladder mucosa with tumor involvement and without new tumor development. Patients who did not achieve a complete or partial remission were considered treatment failures.

Complete responders and those partial responders rendered tumor-free by TUR were offered monthly maintenance therapy at the discretion of the investigator. Patient followup will continue for at least 5 years. At a minimum, medical history, physical examination, and cystoscopy are performed every 3 months with biopsies and/or urine cytologies as indicated.

## RESULTS

Fifty-five of the 63 patients were evaluable for response at the time of this report. An additional two patients, not evaluable for response due to the lack of pathologic evidence of disease prior to therapy, were evaluable for toxicity. The remaining six patients who were entered in the study had not completed therapy.

### Toxicity

Table 2 describes the nature and incidence of reported toxic effects by dose. There were four patients (7%) who experienced lower urinary tract irritation manifested by dysuria and/or urinary frequency. Symptoms cleared without therapy. Three patients (5%) had a transient rash on the hands which cleared spontaneously or with topical corticoid cream. Neither reaction necessitated an interruption of therapy and all but one of these patients achieved a complete remission. No clinical evidence of systemic toxicity was noted. Liver profile studies, renal function tests, and hematologic parameters remained within normal limits. No local reactions occurred at < 30 mg/week.

TABLE 2.—Nature and incidence of toxicity

| Dose (mg/wk) | No. of patients (%) with— | | |
|---|---|---|---|
| | Systemic effects | Bladder irritation | Rash |
| 20 | 0 | 0 | 0 |
| 25 | 0 | 0 | 0 |
| 30 | 0 | 2 (3.5) | 2 (3.5) |
| 40 | 0 | 2 (3.5) | 0 |
| 60 | 0 | 0 | 1 (1.8) |
| Total | 0 | 4 (7) | 3 (5) |

### Absorption and Urine Concentration

The initial concentration of mitomycin C instilled was 1000 or 2000 μg/ml depending on the dose. During the 2-hour retention period, the median urinary efflux (urine plus instilled fluid) was 150 ml (range, 50–400 ml). The concentration of mitomycin C in the efflux ranged from 89 to 400 μg/ml as shown in table 3. No absorption of mitomycin C from the bladder was detected at any dose level based on serial serum sample assays.

### Response

The overall response rate was 67% (table 4). There were approximately twice as many complete (45%) as partial (22%) remissions reported. No statistical differences in response rates occurred between dose groups, and treatment failures occurred with equal frequency. Patients were demographically comparable.

Table 5 describes the differences in disease status by response to therapy. Although this study was not designed to determine the prognostic significance of various disease characteristics, there appeared to be a trend suggesting a negative prognosis for an increased annual recurrence rate as shown by other authors (19). However, the differences in this patient population were not statistically different from the overall median recurrence rate. Prior therapy and extent of disease were comparable among the response groups as was pathologic grade in those patients with adequate medical histories.

Although some partial responders had sufficient reduction of tumor burden to permit TUR of residual tumor, the most significant response was a complete remission. The characteristics of each complete responder are listed in table 6. Remissions of relatively prolonged duration were noted among patients who had received prior chemotherapy with moderate or extensive disease even at the lowest doses. The duration of followup is not yet sufficient to determine the median duration of response for the higher dose groups, but the median duration of complete response is 14.5 months, with a range of 1–36+ months for those patients treated at the two lower doses.

One patient with extensive grade II disease, who had a partial response to prior intravesical bleomycin, was treated with 20 mg/week of mitomycin C. She achieved a complete remission lasting 4 months. At recurrence she received 25 mg of mitomycin C/week and again had a complete remission, this time for 21 months. Similarly, a patient with moderate grade II tumor achieved a partial response to 30 mg of mitomycin C/week and then under-

TABLE 3.—Absorption and urinary concentration* of mitomycin C by dose

| Dose (mg/wk) | No. of evaluable patients | Mitomycin C absorption (No. of patients) | Median urinary concentration (μg/ml) (range) |
|---|---|---|---|
| 20 | 15 | 0 | 89 (50–175) |
| 25 | 7 | 0 | 92 (75–222) |
| 30 | 16 | 0 | 167 (53–290) |
| 40 | 7 | 0 | 153 (39–310) |
| 60 | 3 | 0 | 400 (200–660) |

*Urinary concentration = urine efflux + instilled fluid.

TABLE 4.—Response rate by dose of mitomycin C

| Dose (mg/wk) | No. of evaluable patients/ No. entered | No. of responses (%) | | |
|---|---|---|---|---|
| | | Partial | Complete | Complete + partial |
| 20 | 15/15 | 3 (20) | 6 (40) | 9 (60) |
| 25 | 10/10 | 1 (10) | 6 (60) | 7 (70) |
| 30 | 16/20 | 4 (25) | 6 (38) | 10 (63) |
| 40 | 8/11 | 3 (38) | 3 (38) | 6 (75) |
| 60 | 6/7 | 1 (17) | 4 (67) | 5 (83) |
| Total | 55/63 | 12 (22) | 25 (45) | 37 (67) |

went a TUR of residual tumor. At recurrence, the patient received a 60-mg/week course of mitomycin C and was again rendered tumor-free by a partial response to chemotherapy and TUR. These cases indicate that a recurrence does not preclude a second meaningful response and that a higher dose may induce a longer disease-free interval.

## DISCUSSION

These results demonstrate that mitomycin C can be given at weekly doses of 20–60 mg without significant toxicity. There were no incidents of systemic toxicity reported and no absorption of drug from the bladder at any dose. Local bladder irritation occurred infrequently, did not require any interruption of therapy, and resolved spontaneously. The etiology of the rash occurring in three patients is unknown, but may have been an allergic reaction.

The current 45% complete remission rate includes nine patients (36%) who had failed or had a recurrence after prior intravesical therapy with one or more chemotherapeutic agents. The overall complete response rate compares favorably with the consistently reported one third complete remission rate for thioTEPA therapy in previously untreated patients. The complete responses in nine patients who had received prior intravesical chemotherapy suggest that mitomycin C may be an effective second-line agent as well. Duration of complete remission of > 14 months for patients receiving mitomycin C appears to exceed that of patients receiving thioTEPA (19), although numbers are small in both studies. Partial remission was achieved by 22% of the patients in this study and often resulted in a reduction in tumor burden sufficient to permit complete TUR of residual tumor.

Although no clear dose-response relationships are shown, the total remission rate appears to be increasing at higher doses. The number of patients at each dose is

TABLE 5.—Disease characteristics by response to mitomycin C therapy

| Disease characteristic | Response | | |
|---|---|---|---|
| | Complete (25)* | Partial (12) | Failure (18) |
| Median duration of disease in mos† | 17 (20) | 9 (9) | 16 (14) |
| Median No. of yearly recurrences | 1.7 (20) | 2.2 (9) | 2.5 (14) |
| % of patients with— | | | |
| Prior chemotherapy or radiotherapy | 32 (23) | 27 (12) | 33 (16) |
| Pathologic grade | | | |
| I–II | 92 | 92 | 94 |
| III | 8 | 8 | 6 |
| | (17) | (7) | (7) |
| Extent of disease | | | |
| Limited | 68 | 58 | 72 |
| Moderate | 24 | 25 | 22 |
| Extensive | 8 | 17 | 6 |
| | (21) | (12) | (15) |

*Nos. in parentheses = No. of evaluable patients.
†Time from diagnosis to mitomycin C therapy.

TABLE 6.—Characteristics of complete responders by dose of mitomycin C

| Dose (mg/kg) | Prior chemotherapy* (response) | Extent of disease | Duration of complete response (mos) |
|---|---|---|---|
| 20 | | | |
| | BLM (partial) | Extensive | 4 |
| | TTP (failure) | Moderate | 36+ |
| | — | Limited | 28 |
| | — | Moderate | 8+† |
| | TTP (failure) | Limited | 1 |
| | — | Limited | 2+† |
| 25 | | | |
| | — | Limited | 30+ |
| | MMC (complete) | Extensive | 21 |
| | — | Limited | 27+ |
| | — | Moderate | 21+ |
| | — | Limited | 6 |
| | TTP (complete, partial)<br>BLM (failure) | Moderate | 4+‡ |
| 30 | | | |
| | — | Limited | 30+ |
| | — | Moderate | 3 |
| | — | Limited | 10+ |
| | — | Limited | 9+ |
| | — | Limited | 14+ |
| | DDP (failure) | Limited | 1+† |
| 40 | | | |
| | — | Moderate | 13+ |
| | — | Limited | 6 |
| | TTP (unknown) | Limited | 6+ |
| 60 | | | |
| | — | Limited | 5+ |
| | TTP (unknown) | Limited | 7 |
| | — | Limited | 4+ |
| | — | Limited | 4+ |

*TTP = thioTEPA; BLM = bleomycin; MMC = mitomycin C; and DDP = *cis*-dichlorodiammineplatinum(II).
†Lost to followup.
‡Died with no evidence of disease.

small and includes a heterogeneous group of tumors. That the duration of complete remission may increase with dose is suggested by the patient whose second complete response was five times longer than the first after a higher dose of mitomycin C was administered. Clearly, there was no toxic liability with higher doses. Consequently, we have selected 40 mg/week as the dose of mitomycin C to employ in phase II–III studies. The weekly schedule of drug administration was chosen for patient convenience and these data suggest it to be as active as the three times weekly regimen.

## REFERENCES


1. SCHADE ROK, and SWINNEY J. The association of urothelial atypism with neoplasia: its importance in treatment and prognosis. J Urol 109:619–622, 1973.
2. WELDON TE, and SOLOWAY MS. Susceptibility of urothelium to neoplastic cellular implantation. Urology 5:824–827, 1975.
3. ABBASSIAN A, and WALLACE DM. Intracavitary chemotherapy of diffuse non-infiltrating papillary carcinoma of the bladder. J Urol 96:461–465, 1966.
4. VEENEMA RJ, DEAN AL, JR, USON AC, ET AL. Thiotepa bladder instillations: therapy and prophylaxis for superficial bladder tumors. J Urol 101:711–715, 1969.
5. NATIONAL BLADDER CANCER COLLABORATIVE GROUP A. The role of intravesical thiotepa in the management of superficial bladder cancer. Cancer Res 37:2916–2917, 1977.
6. NIEH PT, DALY JJ, HEANEY JA, ET AL. The effect of intravesical thio-tepa on normal and tumor urothelium. J Urol 119:59–61, 1978.
7. JONES HC, and SWINNEY J. Thio-TEPA in the treatment of tumors of the bladder. Lancet 2:615–618, 1961.
8. Case records of the Massachusetts General Hospital. Weekly clinicopathological exercises. Cases 28-1977. N Engl J Med 297:102–107, 1977.
9. GROSSMAN BH. Current therapy of bladder carcinoma. J Urol 121:1–7, 1979.
10. MISHINA T, ODA K, MURATA S, ET AL. Mitomycin C bladder instillation therapy for bladder tumors. J Urol 114:217–219, 1975.

11. Shida K, Doguchi T, Sasazaki T, et al. Treatment of the bladder cancer with instillations of mitomycin C. Jpn J Clin Urol 21:1075–1079, 1967.
12. Nishiura T, Kumamoto E, Nishimura Y, et al. Effects of mitomycin C intracavitary therapy for recurrent bladder cancers. Ijaku No Ayumi 65:637–640, 1968.
13. Sai E, and Hayakawa T. Intracavitary therapy for bladder tumors using mitomycin. Med Consult New Remedies 5:1933, 1968.
14. Ogawa H. Intracavitary treatment of tumors of the bladder with cytotoxic agent (mitomycin C) and radioactive phosphorus. I. Jpn J Urol 60:717, 1969.
15. Imamura K, Yoshida H, Nakano H, et al. Effects of intravesical treatment of bladder tumors with mitomycin C. Acta Urol Uap 20:33–38, 1974.
16. Bracken RB, and Johnson D. Treatment of multiple superficial tumors of the urinary bladder with intravesical mitomycin C. *In* Mitomycin C—Current Status and New Developments (Carter SK, Umezawa H, and Crooke ST, eds). New York, NY, Academic Press, 1979, pp 205–211.
17. Schwartz HS, Sedergren JE, and Phillips FSS. Mitomycin C: chemical and biological studies on alkylation. Science 142:1181–1183, 1963.
18. Fujita H. Comparative studies on the blood level tissue distribution, excretion and inactivation of anticancer drugs. Jpn J Clin Oncol 12:151–162, 1971.
19. Althausen AF, Prout GR, Jr, and Daly JJ. Non-invasive papillary carcinoma of the bladder associated with carcinoma in situ. J Urol 116:575–580, 1976.

# Combination Antifol Therapy: Observations on the Development of Resistant L1210 Leukemic Cells In Vivo [1]

George P. Browman [2,3,*] and Herbert Lazarus [2]

## SUMMARY

**L1210 murine leukemic cells were serially passaged in $BDF_1$ mice and were treated in vivo with methotrexate (MTX) and 2,4-diamino-5-(3′,4′-dichlorophenyl)-6-methylpyrimidine (DDMP) alone or in combination. The time course of emergence of resistance of the treated cell lines was studied both in vivo and in vitro. When used as a single drug, in vivo resistance to MTX developed gradually and was considerable at eight passages and complete by 11 passages. Complete in vivo resistance to DDMP used alone occurred by the fifth passage. Complete in vivo resistance to the drugs in combination was seen by the eighth passage. In cells demonstrating complete DDMP resistance, as determined in vivo, there was no evidence of cross-resistance to MTX measured either in vivo or in vitro, while MTX resistance was associated with incomplete cross-resistance to DDMP. The greatest degree of resistance, as determined in vitro, occurred in the cell line treated with the drug combination. In vitro tests of drug resistance correlated well with in vivo survival data. An important observation was that major in vivo drug resistance was accompanied by only a small measurable effect using standard in vitro screening techniques. The implications of this finding are discussed.**

[Cancer Treat Rep 64:231–236, 1980]

Methotrexate (MTX) and 2,4-diamino-5-(3′,4′-dichlorophenyl)-6-methylpyrimidine (DDMP) are both antifols which are active in experimental animal tumor systems and in human cancer. The antifolate effect of MTX is due to inhibition of the enzyme, dihydrofolate reductase, resulting in depletion of reduced folate pools. There is strong evidence that DDMP exerts its antifolate effect in the same manner (1,2). These compounds differ substantially in other respects. Structurally, MTX is a folate analog while DDMP possesses a substituted pyrimidine ring. These compounds also have quite different pharmacologic properties which in large part are due to differences in lipid solubility (3). These pharmacologic differences, at least partially, account for the effectiveness of DDMP in experimental tumors which are resistant to MTX (4).

MTX has found clinical usefulness in the treatment of a wide variety of tumors. Its therapeutic benefits, however, are partially limited by the development of tumor resistance and by its relative insolubility at acid pH which limits the dose which can be safely administered in rescue protocols (5,6). The clinical usefulness of DDMP has been limited by neurotoxic manifestations most likely related to inhibition of the enzyme histamine, *N*-methyl transferase, resulting in elevated levels of histamine (3).

In view of the established value of antifols in the treatment of some human cancers, the efficacy of rescue protocols to prevent toxicity, and the widely differing non-antifol-related toxic effects of MTX and DDMP, the effectiveness of these compounds in combination was tested. The present studies examine the patterns of resistance which develop when MTX and DDMP are used in combination on an experimental tumor system.

[1]Received June 11, 1979; revised Nov 26, 1979; accepted Dec 5, 1979.

[2]Sidney Farber Cancer Institute, and Departments of Medicine and Pathology, Harvard Medical School, Boston, MA.

[3]G. P. Browman is a recipient of a fellowship from the Medical Research Council of Canada.

**Reprint requests to*: George P. Browman, MD, Ontario Cancer Treatment and Research Foundation, Hamilton Centre, 711 Concession St, Hamilton, Ontario L8V 1C3, Canada.

## MATERIALS AND METHODS

### In Vivo Studies

L1210 murine leukemic cells and male $BDF_1$ (C57BL × $DBA_2$) mice were obtained from the Arthur D. Little Co (Cambridge, MA). MTX was supplied by the Division of Cancer Treatment of the National Cancer Institute (Bethesda, MD) as a lyophilized powder and was reconstituted with water prior to use. DDMP was a gift from Dr. C. A. Nichol, Burroughs-Wellcome Co (Research Triangle Park, NC), and was suspended in a solution of 10% absolute alcohol, 10% Emulphor, and 80% normal saline prior to injection.

Animals were divided into three treatment groups that received either MTX alone, DDMP alone, or MTX-DDMP combined, according to the regimens shown in table 1. An untreated group was also included. On Day 0, $10^6$ L1210 cells were injected ip and drug treatment was started on Day 1. At about 7 days, when ascites was evident, cells were harvested from the peritoneal cavity by aspiration with a Pasteur pipette under sterile conditions. Cells were washed three times in prewarmed unsupplemented Hanks' medium and were resuspended at $2 \times 10^6$ cells/ml, 0.5 ml of which was injected ip into fresh mice that received drug treatment 1 day later. Cells were serially passaged in this manner and an aliquot of cells from each passage was cryopreserved in fully supplemented Eagle's minimal essential medium in 10% dimethylsulfoxide in an LR40 liquid nitrogen freezer (Linde Corp) for future in vivo and in vitro studies.

Cell lines from specific passages were selected for screening of in vivo drug resistance by injection into groups of mice (ten per group) for determination of median survival time (MST). Resistance to a drug was determined by comparing the MST of a group of drug-treated animals to that of an untreated group bearing the same tumor and then calculating the percent increase in lifespan (%ILS) of the treated group according to the formula: $\%ILS = \left(\frac{MST\ (treated)}{MST\ (untreated)} - 1\right) \times 100.$

In vivo screening involved the same drug dose schedules used for development of resistance in serial passages (table 1).

### In Vitro Studies

For routine passage, and during dose-response experiments, cells were maintained in asynchronous logarithmic growth at 37°C in Eagle's minimal essential medium (Microbiological Associates, Bethesda, MD) supplemented with 10% fetal bovine serum, $1 \times 10^{-5}$ M 2-mercaptoethanol, streptomycin sulfate (100 μg/ml), sodium penicillin G (100 units/ml), Hepes buffer (20 mM), and L-glutamine (0.03%).

For in vitro dose-response curves, DDMP was dissolved in lactate and the pH was adjusted with NaOH so that addition to cell cultures would not result in a pH change. DDMP and MTX were diluted in unsupplemented medium and added to cell cultures in 1:100 dilution to achieve the desired drug concentrations. Cell cultures were set up at $10^5$/ml in triplicate for each drug concentration in multiwell plates, and after 48 hours of continuous drug exposure they were harvested and counted using a Coulter Counter (Model F). A control untreated set of cultures was included for each separate dose-response experiment. From the data obtained, a dose-response curve was drawn and the ID50 was calculated according to the method described by Jackson et al (7). The ID50 represents the concentration of drug that halves the growth rate relative to untreated controls. The degree of in vitro resistance was expressed as the ratio of the ID50 values of cells exposed to drug in vivo versus cells which had been passaged six times in vivo without previous drug exposure (control cell line).

## RESULTS

### In Vivo Studies

The antitumor effect obtained by the different treatment regimens on previously untreated serially passaged L1210 cells is shown in table 2. The %ILS produced by these regimens was constant through 12 passages (data not shown), and for subsequent experiments the sixth passage of untreated cells was used as the control line. MTX alone and the DDMP-MTX combination resulted in an equivalent prolongation of survival rate (59% and 63%) while DDMP alone resulted in a slightly longer sur-

TABLE 1.—Treatment regimens employed for development of drug resistance in L1210 murine leukemia in vivo

| Drug regimen | Route | Schedule |
|---|---|---|
| DDMP alone | sc | 40 mg/kg, Day 1 |
| MTX alone | ip | 3.0 mg/kg, Days 1–4 |
| Combination | | |
| DDMP | sc | 20 mg/kg, Day 1 |
| and | | and |
| MTX | ip | 1.5 mg/kg, Days 1–4 |

TABLE 2.—Effect of treatment regimens on survival of mice bearing control L1210 cells previously unexposed to drug*

| Drug regimen | MST (%ILS ± SE) |
|---|---|
| DDMP alone | 79 ± 8(3)† |
| MTX alone | 59 ± 3(6) |
| Combination DDMP and MTX | 63(1) |

*Antitumor effect remained unchanged through 12 passages (data not shown).
†Nos. in parentheses = No. of experiments performed.

vival rate (79%). The effect of multiple courses of treatment on the survival of tumor-bearing mice is shown in table 3. Cells treated with MTX alone demonstrated considerable resistance to the drug by the eighth passage, and they were completely resistant to the drug dose administered by the 11th passage. Cells treated with DDMP were completely resistant to the administered dose by the fifth passage. When these compounds were used in combination, complete resistance to the administered dose of both agents occurred simultaneously at the eighth passage. The pattern of cross-resistance which developed during MTX and DDMP treatment is also shown in table 3. By the fifth passage, there was partial resistance of MTX-treated cells to MTX and this was accompanied by a minor degree of DDMP cross-resistance in vivo. By the eighth passage, the degree of resistance to DDMP was equal to that of MTX at the drug doses used. Despite the early development of complete resistance to the DDMP dose used, no significant in vivo cross-resistance to MTX was seen through eight passages.

## In Vitro Studies

Once in vivo resistance to MTX and DDMP was established, the cells were grown in vitro and their sensitivity to DDMP and MTX was determined using dose-response experiments and calculating an ID50 value for each drug. Major resistance to the in vivo dose schedule of MTX (passage 8) was associated with an in vitro ID50 ratio of the MTX-treated tumor line to the control tumor cell line (untreated cells at passage 6) of 2.7 (table 4). For the DDMP-treated tumor line, complete resistance to the in vivo dose schedule used (passages 5 and 7) was associated with an in vitro ID50 ratio of 4.8. The greatest degree of in vitro resistance to MTX and DDMP was seen in the cells exposed in vivo to the combination schedule with an in vitro ID50 ratio of 18.1 (MTX) and 7.1 (DDMP) at the 11th passage.

The in vitro data further confirmed the in vivo cross-resistance patterns (table 5). Cell lines made resistant to

TABLE 3.—Effect of multiple courses of drug treatment on tumor response and development of cross-resistance patterns through serial passages

| In vivo drug exposure | Passage No. | In vivo drug screening: MST (%ILS ± SE) Resistance | Cross-resistance | |
|---|---|---|---|---|
| | | *MTX* | | *DDMP* |
| L1210 MTX | 0 | 59 ± 3(6)* | | |
| | 3 | 38(1)† | | |
| | 5 | 29,25(2)† | | 50(1) |
| | 8 | 20 ± 11(3) | | 22,25(2) |
| | 11 | 0(1) | | 22(1) |
| | | *DDMP* | | *MTX* |
| L1210 DDMP | 0 | 79 ± 8(3) | | |
| | 1 | 25(1) | | |
| | 3 | 14,14(2) | | 43,43(2) |
| | 5 | 0,0(2) | | 43,38(2) |
| | 7 | 0(1) | | |
| | 8 | | | 43(1) |
| | | *MTX* | *DDMP* | *MTX + DDMP* |
| L1210 MTX + DDMP | 0 | 59 ± 3(6) | 79 ± 8(3) | 63(1) |
| | 3 | | | 38(1) |
| | 5 | 43(1) | 71(1) | |
| | 6 | 12(1) | 50(1) | |
| | 8 | 4 ± 4(3) | 4 ± 4(3) | 0(1) |

*Nos. in parentheses = No. of separate experiments performed.
†Where < 3 separate experiments performed all data are shown.

TABLE 4.—Correlation of in vivo and in vitro expressions of resistance

| In vivo drug exposure | Passage No. | In vivo resistance (mean %ILS ± SE) | | In vitro resistance (mean ID50 $\frac{\text{treated cells}}{\text{control cells}}$ ± SE) | |
|---|---|---|---|---|---|
| | | *MTX* | | *MTX* | |
| L1210 MTX | 1 | | | 0.96(1) | |
| | 5 | 29,25(2)*,† | | 1.5,1.5(2) | |
| | 8 | 20 ± 11(3) | | 2.7 ± 0.1(3) | |
| | | *DDMP* | | *DDMP* | |
| L1210 DDMP | 1 | 25(1) | | 1.3(1) | |
| | 3 | 14,14(2) | | 3.6,5.7(2) | |
| | 5+7 | 0(3) | | 4.8 ± 0.5(7) | |
| | | *MTX* | *DDMP* | *MTX* | *DDMP* |
| L1210 MTX + DDMP | 2 | | | 1.2(1) | |
| | 5 | 43(1) | 71(1) | 1.7 ± 0.1(3) | 0.92,0.93(2) |
| | 8 | 4 ± 4(3) | 4 ± 4(3) | 6.8 ± 0.4(3) | 5.8 ± 0.5(3) |
| | 11 | | | 18.1(1) | 7.1(1) |

*Nos. in parentheses = No. of separate experiments performed.
†Where < 3 experiments performed all data are shown.

TABLE 5.—Correlation of in vivo and in vitro cross-resistant patterns

| Cell line | Passage No. | In vivo (mean %ILS ± SE) | | In vitro (mean ID50 $\frac{\text{treated cells}}{\text{control cells}}$ ± SE) | |
|---|---|---|---|---|---|
| | | Resistance | Cross-resistance | Resistance | Cross-resistance |
| L1210 MTX | 8 | 20 ± 11(3)* | 22,25†(2) | 2.7 ± 0.1(3) | 4.3 ± 0.6(3) |
| L1210 DDMP | 5 + 7 | 0(3) | 37 ± 4(4) | 4.8 ± 0.5(7) | 1.00 ± 0.07(6) |

*Nos. in parentheses = No. of experiments performed.
†Where < 3 experiments performed all data are shown.

MTX in vivo (at passage 8) and demonstrating some in vivo DDMP cross-resistance had an in vitro ID50 ratio of 4.3 for DDMP. The cell lines made resistant to DDMP in vivo (at passages 5 and 7) were not cross-resistant to MTX in vivo and showed an in vitro ID50 ratio of 1.00 for MTX.

The in vitro ID50 ratios also reflected development of in vivo resistance during successive tumor passages. This is illustrated in the cell line treated with the combination (table 4). At passage 5, there was minor resistance to MTX (ILS, 43%), and no evidence of resistance to DDMP (ILS, 71%). The corresponding in vitro ID50 ratios were 1.7 for MTX and 0.93 for DDMP. By passage 8, almost complete in vivo resistance had developed to the administered dose schedules of both agents, and the in vitro ID50 ratios correspondingly increased.

When the cells were exposed to the drug combination through three further passages there was a further increase of the in vitro ID50 ratio for both drugs. At the early passages for both MTX- and combination-treated cell lines, when no resistance could be demonstrated in vivo, the in vitro ID50 ratio was 1.04 for MTX.

## DISCUSSION

Antifols, in particular MTX, have proven to be extremely useful in treating a wide variety of human malignancies. The relatively recent introduction of high-dose MTX therapy has shown that, in some cases, "resistant" tumors can be made to respond (8). The drug doses which can be utilized, however, are limited by toxic effects which are not directly related to their antifolate action (eg, renal toxicity). In clinical trials, DDMP has demonstrated antitumor activity, but treatment has been restricted by the occurrence of neurotoxic manifestations (9–11). Due to these considerations, it was felt that combinations of antifols which possess different limiting toxic effects might be useful. MTX and DDMP were used in combination in vivo and the effect of this combination on the development of resistance was studied and compared with the resistance developed by the use of each agent alone. In vivo and in vitro measurements of drug resistance were correlated. Three important observations emerged from the results. First, the use of antifols in combination may delay onset of resistance; second, re-

sistance to an antifol such as DDMP is not necessarily associated with MTX cross-resistance; and third, an easily measurable degree of in vivo resistance may be difficult to prove with current in vitro methods.

At the drug doses used, resistance developed more quickly to DDMP than to MTX. The use of the combination resulted in a significant delay in the onset of resistance to DDMP, but once antifol resistance did develop, it was greatest with the two-drug combination.

The cross-resistant patterns which developed were not expected. Significant in vivo DDMP resistance was not associated with cross-resistance to MTX measured either in vivo or in vitro. It is obviously important to continue studying the mechanisms of resistance in this line. MTX-resistant cells did show some degree of cross-resistance to DDMP. With combination therapy, once in vivo resistance was apparent, it was associated with resistance to both compounds. Studies of acquired mechanisms of resistance which develop during in vitro exposure of L1210 cell lines to MTX indicate that transport mutations occur early in the selection process and are somewhat unstable, tending subsequently to be replaced by a cell line with elevated dihydrofolate reductase (DHFR) (12). The lack of cross-resistance to MTX in the DDMP-resistant cell line suggests a possible transport defect since a rise in DHFR levels would be expected to confer resistance to both compounds. An alternative explanation could be altered kinetic binding of DDMP for DHFR without an associated change in MTX kinetic binding. Acquired resistance to MTX by this mechanism has been observed (13). Studies are in progress to determine whether DDMP resistance is associated with altered transport of the drug, or changes in DHFR kinetic binding to the inhibitors or the substrate, $FH_2$.

It must be pointed out that the rate at which drug resistance occurs may be a function of the cell-kill achieved by the drug since selection of pre-existing, genetically stable clones is a probable route by which resistance develops. Thus, it may be that the rapid onset of resistance to DDMP in this study simply represents the greater antitumor effect produced by this regimen. It is doubtful, however, that the degree of difference in cell-kill produced by the different regimens accounts for the observed differences in rate of onset of resistance. Furthermore, preliminary evidence suggests that the mechanisms of resistance among the cell lines are different, and this is not explained by different cytotoxic capacities.

An important aspect to emerge from this study concerns the correlative in vitro data and the observation that major acquired in vivo resistance of the tumor is accompanied by relatively small changes in cellular resistance as measured in vitro. This strongly suggests that in the in vivo situation, relatively minor changes in the tumor cells may result in "absolute" clinical resistance. If this is indeed true, then antifol resistance studies using cell lines which are orders of magnitude more resistant than the parent line may bear little relevance to the clinical situation. Furthermore, this observation may explain the difficulties in documenting mechanisms of tumor resistance in man (14).

It may be naive to expect that mechanisms of resistance to certain drugs in human tumors would be the same as those found in experimental tumor lines where the resistant property has been purposely exaggerated. Indeed, it has been suggested for one group of drugs (6-thiopurines) that the degree of resistance may determine the mechanism involved (15). Therefore, mechanisms of resistance which develop to various drugs may be schedule- and dose-dependent and in the early stages our experimental techniques may be too insensitive to record the minor changes which take place. On the other hand, by assuming that mechanisms of drug resistance in human tumors are the same as those induced in experimental tumors, we may be restricting our investigative efforts. The realization that in vivo tumor resistance to drugs may be associated with relatively minor cellular or extracellular changes should broaden our approach in searching for mechanisms of resistance in man. It may be more appropriate to study mechanisms of resistance in cell lines similar to those produced in the present study.

## REFERENCES


1. Bertino J, Iannotti AT, Perkins JP, et al. Dihydrofolate reductase from guinea pig liver and small intestine. Biochem Pharmacol 15:563–571, 1966.
2. Ganeshaguru K, and Hoffbrand AV. Effects of DDMP and DDEP on the deoxyribonucleoside triphosphate concentrations in human cells. Biochem Pharmacol 26:543–545, 1977.
3. Duch DS, Bowers SW, Brenckman WD, Jr, et al. Inhibition of brain histamine-*N*-methyl transferase (HMT) by diaminopyrimidine anticancer agents. The Pharmacologist 18:205, 1976.
4. Nichol CA. Studies on dihydrofolate reductase related to the drug sensitivity of microbiologic and neoplastic cells. Adv Enzyme Regul 6:305–322, 1968.
5. Jacobs SA, Stoller RG, Chabner BA, et al. 7-Hydroxymethotrexate as a urinary metabolite in human subjects and rhesus monkeys receiving high dose methotrexate. J Clin Invest 57:534–538, 1976.
6. Jaffe N, and Traggis D. Toxicity of high-dose methotrexate (NSC-740) and citrovorum factor (NSC-3590) in osteogenic sarcoma. Cancer Chemother Rep Part 3, vol 6(1):31–36, 1975.
7. Jackson RC, Taylor GA, and Harrap KR. Aspects of the biochemical pharmacology of cytembena. Neoplasma 22:259–268, 1975.
8. Djerassi I, Farber S, Abir E, et al. Continuous infusion of methotrexate in children with acute leukemia. Cancer 20:233–242, 1967.
9. Price LA, Goldie JH, and Hill BT. Methodichlorophen as antitumour drug. Br Med J 2:20–21, 1975.
10. Alberto P, Peytremann R, Medenica R, et al. Clinical

experience with a combination of DDMP, a diaminopyrimidine, and folinic acid (CF). Proc Am Assoc Cancer Res and ASCO 19:346, 1978.

11. Miller DS, Rundles RW, Nichol CA, et al. Phase I/II experience with a lipid-soluble folate antagonist: 2,4-diamino-5-(3',4'-dichlorophenyl)-6-methylpyrimidine (DDMP). Proc Am Assoc Cancer Res and ASCO 17:263, 1976.
12. Jackson RC, Niethammer D, and Huennekens FM. Enzymic and transport mechanisms of amethopterin resistance in L1210 mouse leukemia cells. Cancer Biochem Biophys 1:151, 1975.
13. Jackson RC, and Niethammer D. Acquired methotrexate resistance in lymphoblasts resulting from altered kinetic properties of dihydrofolate reductase. Eur J Cancer 13:567, 1977.
14. Bertino JR, and Skeel RT. Resistance to folic acid antagonists—clinical aspects. Biochem Pharmacol 23 (suppl 2):101–106, 1974.
15. Brockman RW. A mechanism of resistance of 6 mercaptopurine: metabolism of hypoxanthine and 6 mercaptopurine by sensitive and resistant neoplasms. Cancer Res 20:643–653, 1960.

# Modified Procarbazine, CCNU, and Vincristine (PCV 3) Combination Chemotherapy in the Treatment of Malignant Brain Tumors [1,2]

**Victor A. Levin,[3,*] Michael S. Edwards,[3] Donald C. Wright,[3] Margaret L. Seager,[3] Tana Pischer Schimberg,[3] Jeanette J. Townsend,[4] and Charles B. Wilson [3,5]**

## SUMMARY

**Fifty-eight patients harboring recurrent malignant brain tumors were evaluated. CCNU (110 mg/m$^2$) was administered on Day 1, procarbazine (60 mg/m$^2$) was administered daily for 14 days beginning on Day 8, and vincristine (1.4 mg/m$^2$) was administered on Days 8 and 29 of each 6-week cycle of therapy. Therapy was continued until tumor progression was documented. Before each course, a neurologic examination was performed and radionuclide and computerized tomographic scans were obtained. Response and progression were defined as improvement or deterioration, respectively, in at least two of the three tests. Of the 46 patients harboring recurrent malignant gliomas, 12 (26%) responded to therapy, 18 (39%) had tumor progression, and 16 (35%) had disease stability. Nineteen of the 46 patients were not previously treated with another form of chemotherapy: eight (42%) responded to therapy, eight (42%) had disease stability, and only three (16%) had early tumor progression. The median time to tumor progression was 26 weeks for patients responding to therapy or having disease stability. Approximately 30% of the patients were alive without evidence of tumor progression at 1 year. Evaluated by time to tumor progression and rate of tumor response or stabilization, this combination was similar to the BCNU-5-fluorouracil combination used for patients harboring recurrent malignant glioma.**

[Cancer Treat Rep 64:237–241, 1980]

This study summarizes our experience with the combination of procarbazine, CCNU, and vincristine (PCV) for the treatment of malignant brain tumors since our previous publication of this drug combination (1). The dose and scheduling of CCNU and procarbazine are different than reported earlier. To differentiate the two studies, we will refer to the earlier protocol as PCV 1 and to the current therapy schedule as PCV 3.

Experiments conducted in our laboratories (2–4) suggested that the administration of vincristine, a cell cycle-specific agent, following the substantial log cell-kill caused by CCNU and procarbazine, might produce a maximum cell-kill during a period when tumor cells should be in a proliferative phase. Extrapolation of the cell cycle kinetics determined for animal tumor cells in vivo (5) gave us a basis for estimating the cell kinetics of human gliomas. Based on these assumptions, our PCV 1 protocol was modified by giving vincristine 7 days after procarbazine. Animal studies (2) also suggested that the systemic toxicity of the PCV 1 protocol could be reduced if procarbazine was administered 7 days after CCNU.

In addition to differences in drug dose and scheduling, PCV 3 therapy was administered to many patients who had received prior chemotherapy, while PCV 1 was evaluated in patients who had shown tumor progression sub-

[1]Received June 22, 1979; revised Oct 24, 1979; accepted Nov 1, 1979.

[2]Supported by Public Health Service grant CA-13525 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare, and Faculty Research Award FRA-155 (*V. A. Levin*) from the American Cancer Society.

[3]Brain Tumor Research Center, Department of Neurological Surgery, School of Medicine, University of California, San Francisco.

[4]Department of Pathology (Neuropathology), School of Medicine, University of California, San Francisco.

[5]We acknowledge the editorial assistance of Neil Buckley.

**Reprint requests to:* Victor A. Levin, MD, Brain Tumor Research Center, Department of Neurological Surgery, School of Medicine, 783 HSW, University of California, San Francisco, CA 94143.

sequent to radiation therapy and who had not received chemotherapy.

## CLINICAL MATERIAL AND METHODS

*Patient selection.*—Patients were accrued sequentially from the Brain Tumor Chemotherapy Service, Department of Neurological Surgery, University of California, San Francisco. All patients had recurrent malignant brain tumors documented by angiogram, radionuclide brain scan, computerized tomographic (CT) brain scan, and clinical neurologic examination. The eligibility requirements were unequivocal deterioration ascribable to tumor regrowth, no surgery for at least 2 months, no radiation therapy for at least 3 months, no chemotherapy for at least 6 weeks, and normal results of cbc, platelet count, SMA-12 panel, urinalysis, and prothrombin time tests. Informed consent was obtained from each patient (and/or responsible next-of-kin) before therapy was initiated.

*Therapy.*—CCNU was administered at a dose of 110 mg/m² orally on Day 1 followed by procarbazine at a dose of 60 mg/m² orally for 14 days beginning on Day 8 of each cycle. Vincristine was administered at a dose of 1.4 mg/m² iv on Days 8 and 29 of each cycle. A full course of therapy was repeated every 6–8 weeks. Toxicity was monitored carefully. CCNU and procarbazine doses were reduced in patients experiencing bone marrow toxic effects (eg, wbcs $\leq$ 3500/mm³ and platelets $\leq$ 100,000/mm³) and were escalated in later courses to levels that could be tolerated by the patient. Vincristine was stopped for severe neurotoxicity (eg, painful paresthesias or peripheral weakness).

Glucocorticoid doses were adjusted as indicated to improve the level of neurologic function and to reduce headache and nausea. Because improvement caused by steroids can mimic response to cytotoxic therapy, doses were increased only when they were required because of marked clinical deterioration. Attempts were made to reduce the steroid dose every 6–8 weeks if the patient's condition was stable or improving. Anticonvulsants and antiemetics were used when medically indicated.

*Evaluation of therapeutic response.*—Effects of this combination therapy were evaluated from results of a neurologic examination and radionuclide and CT scans (6). These tests were performed every 6–8 weeks, and the results were compared to the results of the immediately preceding tests.

Response was defined as an improvement in at least two of the three tests, provided that the patient was receiving a stable or decreasing dose of steroid. Progression was defined as deterioration in at least two of the three tests, provided that the patient had not had his/her dose of steroid decreased since the last course of therapy. Stable disease was defined as no change in the three tests, or improvement in only one test, or deterioration in only one test. Again, the patient had to be receiving a stable or decreasing dose of steroid.

## RESULTS

Table 1 summarizes patient evaluability by tumor type. Forty-six patients received at least one course of PCV 3 therapy; 20 patients did not receive a second course of therapy. Seven of 20 patients died before the second course of therapy; they had a mean Karnofsky performance score of 56 ± 16 (SD) at the time of first treatment. After initiating PCV 3, the median survival time of these patients was 4.5 weeks. The remaining patients (13 of 20) had a mean Karnofsky performance score of 59 ± 11 (SD) when first treated with PCV 3, and had a median survival time of 11 weeks. These 20 patients progressed very rapidly to a moribund state and further therapy was declined by either the patients or their families.

The 46 patients with supratentorial glioblastoma multiforme (GM) and other nonglioblastoma malignant gliomas ([NGM]: anaplastic astrocytoma, malignant glioma, gem-

TABLE 1.—Evaluable and nonevaluable patients by tumor type

| Tumor | No. of patients | | | | |
|---|---|---|---|---|---|
| | Prior chemotherapy | | No prior chemotherapy | | |
| | Fully evaluable | Partially evaluable | Fully evaluable | Partially evaluable | Total |
| Glioblastoma | 5 | 6 (3)* | 1 | 0 | 12 |
| Anaplastic astrocytoma | 1 | 6 (3) | 4 | 1 (1) | 12 |
| Malignant glioma | 4 | 3 | 2 | 2 | 11 |
| Gemistocytic astrocytoma | 1 | 0 | 1 | 0 | 2 |
| Mixed malignant astrocytoma | 1 | 0 | 3 | 0 | 4 |
| "Low-grade" astrocytoma | 0 | 0 | 5 | 0 | 5 |
| Total | 12 | 15 (6) | 16 | 3 (1) | 46 |

*Nos. in parentheses = No. of patients who died before receiving a 2nd course of PCV 3.

TABLE 2.—Evaluation of response for all GM and NGM patients including "partially evaluable" patients

| Tumor type | Responders* | | Stable disease* | | Disease progression* | |
|---|---|---|---|---|---|---|
| | PC† | NPC‡ | PC | NPC | PC | NPC |
| GM | 2/11 | 0/1 | 3/11 | 1/1 | 6/11 | 0/1 |
| NGM | 2/16 | 8/18 | 5/16 | 7/18 | 9/16 | 3/18 |
| Subtotal (%) | 4/27 (15) | 8/19 (42) | 8/27 (30) | 8/19 (42) | 15/27 (56) | 3/19 (16) |
| Total (%) | 12/46 (26) | | 16/46 (35) | | 18/46 (39) | |

*No. of patients/total No. evaluable.
†PC = received chemotherapy before PCV 3.
‡NPC = did not receive chemotherapy before PCV 3.

istocytic astrocytoma, and mixed malignant glioma) were analyzed separately and form the basis of this study. Data on all GM and NGM patients were analyzed for response, which assumed all patients were evaluable if they received only one course of chemotherapy (table 2). Patients who had tumor progression with prior chemotherapy (usually BCNU or thioTEPA) administered subsequent to radiation therapy had poor response or disease stabilization to PCV 3 (15% of patients with response and 30% with disease stabilization) compared with patients who had received no prior chemotherapy (42% of patients with response and 42% with disease stabilization).

The second method of analysis was based on the 28 patients who received at least two courses of PCV 3 therapy. Figure 1 shows the Kaplan-Meier (7) representations for weeks to tumor progression from the first course of PCV 3 for patients with GM who had not received prior chemotherapy, for patients with NGM who had received prior chemotherapy, and for patients with NGM who had not received prior chemotherapy. While there are only five, 15, and eight patients in the three groups, respectively, the data are consistent with previous studies in that they show median times to progression of 14, 27, and 26 weeks for the three groups, respectively.

Even though patients with GM generally do worse than patients with NGM, we pooled both groups and analyzed time to tumor progression as a function of Karnofsky performance score at the time of first treatment with PCV 3 therapy. Figure 2 is a Kaplan-Meier representation of time to progression for patients with Karnofsky performance scores of $\geq 70$. The modest differences in time to progression, which suggest that patients who had no prior chemotherapy do better, were not significant by the Gehan test for censored data (8).

Figure 3 shows the average CCNU and procarbazine dose (milligrams per meter squared per week per course) of PCV 3 for patients with response and stable disease. Patients with prior chemotherapy are compared with those without prior chemotherapy. As would be expected, patients receiving prior chemotherapy with other agents generally tolerated less CCNU and procarbazine.

## DISCUSSION

While PCV 3 showed clear-cut activity against recurrent gliomas, it had greater activity in patients who had not previously received chemotherapy. Of patients who had tumor recurrence subsequent to radiation therapy alone (no chemotherapy), 42% responded to therapy and 42% had disease stability, whereas of patients who had failed prior chemotherapy, only 15% responded and 30% had disease stability. However, if patients who had re-

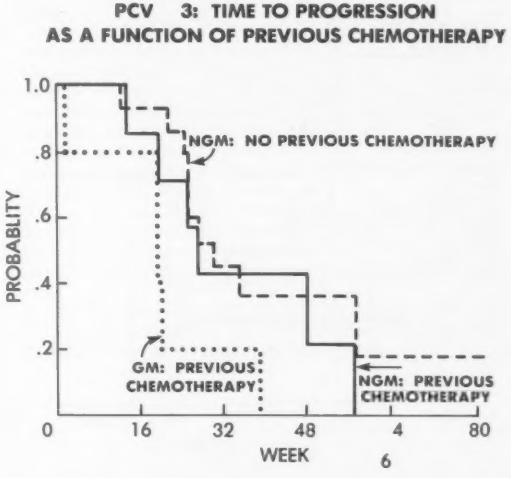



FIGURE 1.—Kaplan-Meier representation of time to tumor progression for patients (5) with GM who received prior chemotherapy (. . . .) and for patients with NGM who were treated for tumor recurrence following radiation therapy alone (— — —) or following radiation therapy and prior chemotherapy (————).

sponse or disease stabilization to PCV 3 were evaluated separately, no statistically significant difference in time to tumor progression was found for patients with NGM (fig 1) or for patients with GM and NGM who had Karnofsky performance scores of ≥ 70 at entry in the study (fig 2). Median times to progression of 26–29 weeks were found for these groups.

Because of the number of phase II trials we have conducted, we compared the results of PCV 3 with the results of PCV 1 and the results of our BCNU-5-fluorouracil (5-FU) study (1,9). Retrospective comparisons of cancer therapy trials are difficult, even within the same institution. While patient referral patterns, delivery of patient care, quality of neurodiagnostic tools, and other factors affecting therapy are constantly changing, PCV 3 and BCNU-5-FU phase II trials were conducted at approximately the same time and are readily comparable. Comparison of PCV 1, a study that ended in September 1973, with PCV 3 was more difficult and less valid because of differences in neurodiagnostic evaluation by CT scanning.

Criteria for tumor progression with PCV 1 therapy were unequivocal deterioration in the neurologic examination *and* radionuclide brain scan. For the PCV 3 and BCNU-5-FU studies, the CT scan was included in the serial evaluations (6). Our current criteria are unequivocal deterioration in two of the three tests, provided that steroid doses are stable or decreasing. These new criteria affect the evaluation of progression for patients with tumors in "silent" areas of the brain such as the frontal lobe, right temporal lobe, and thalamus. Patients in the PCV 3 and BCNU-5-FU studies would be expected to have tumor progression detected earlier than patients in the PCV 1 study, and we estimate that there is a 6-week difference in time to tumor progression for 20%–30% of the patients in the more recent studies.

Figure 3 suggests that, in terms of dose per week, patients receiving PCV 1 tolerated less CCNU after the

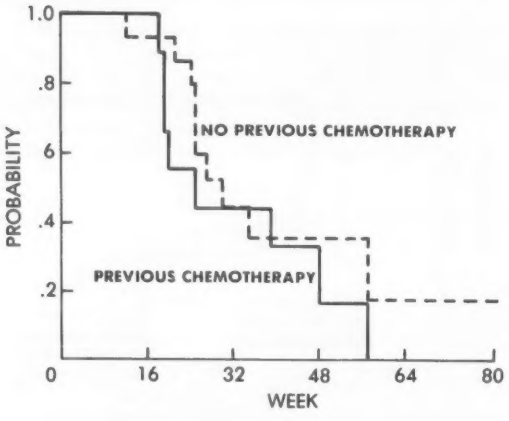



FIGURE 2.—Kaplan-Meier representation of time to tumor progression for patients with NGM and GM, with Karnofsky performance scores of ≥ 70 at first treatment. Patients with no previous exposure to chemotherapy (— — —) are compared to those with prior chemotherapy (————). No difference was found between the 2 groups.

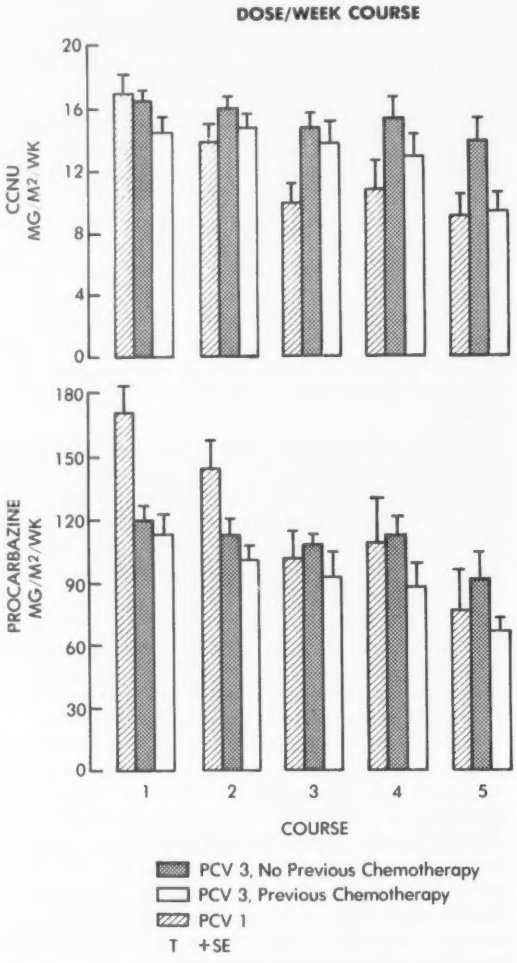



FIGURE 3.—Average CCNU and procarbazine dose (mg/m²/wk/course) of therapy for the 1st 5 courses of PCV 1 and PCV 3. Patients who had had chemotherapy before PCV 3 tolerated less CCNU and procarbazine than patients who had not had chemotherapy previously. The differences increase with each course and are significant ($P < 0.01$).

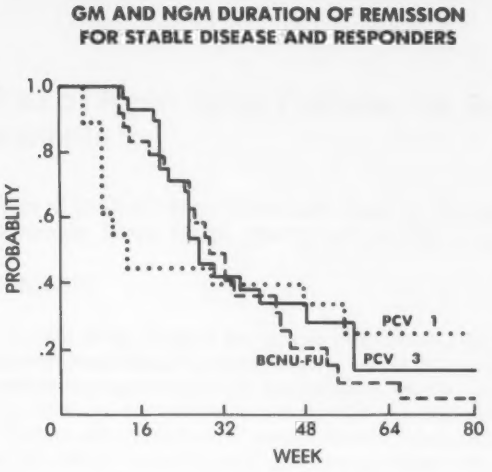



FIGURE 4.—Kaplan-Meier representation of time to tumor progression for patients showing response to therapy or disease stabilization (ie, no progression). PCV 3 (————), BCNU-5-FU (FU) (— — —), and PCV 1 (. . . .) are compared. PCV 1 was worse than the other 2 treatment regimens with a $P < 0.06$ (Gehan test, ref 8).

third course but tolerated more procarbazine for the first two courses. It is likely that scheduling rather than dose may be the important differentiating feature of these two treatments.

A second problem in comparing PCV 1 with PCV 3 is that only one patient in the PCV 1 study group had received prior chemotherapy, while 59% of patients in the PCV 3 trial had received prior chemotherapy (BCNU, thioTEPA, and VM-26). Therefore, we first compared patients who had not received prior chemotherapy (patients who showed recurrence subsequent to radiation therapy). Approximately 40% of patients receiving PCV 1 and PCV 3 responded to therapy; 42% of patients receiving PCV 3 were stable, and only 20% of patients receiving PCV 1 were stable.

When PCV 3 was compared with BCNU-5-FU (9), our current "best therapy" for patients harboring recurrent malignant gliomas, the percentage of patients demonstrating response and/or stabilization of disease progression was the same (84% versus 83%). When time to tumor progression was compared for all three studies for patients with disease stability or response to therapy, PCV 3 was almost identical to BCNU-5-FU and both were slightly better than PCV 1 ($P = 0.056$), particularly in the first 10 weeks of treatment (fig 4).

In summary, PCV 3 produces disease remission and a duration of remission with an efficacy almost identical to BCNU-5-FU. PCV 3 appears to be somewhat superior to PCV 1 in the number and duration of tumor remissions, particularly when the difference in evaluation is considered: with current criteria for tumor progression, a significant number of patients in our original report (1) would have shown progression 6 weeks earlier. Disappointingly, we have not yet devised therapy regimens that can produce and maintain long-term remission.

## REFERENCES

1. GUTIN PH, WILSON CB, KUMAR AR, ET AL. Phase II study of procarbazine, CCNU, and vincristine combination chemotherapy in the treatment of malignant brain tumors. Cancer 35:1398–1404, 1975.
2. LEVIN VA, and WILSON CB. Correlations between experimental chemotherapy in the murine glioma and effectiveness of clinical therapy regimens. Cancer Chemother Pharmacol 1:41–48, 1978.
3. ROSENBLUM ML, KNEBEL KD, VASQUEZ DA, ET AL. In vivo clonogenic tumor cell kinetics following 1,3-bis(2-chloroethyl)-1-nitrosourea brain tumor therapy. Cancer Res 36:3718–3725, 1976.
4. NOMURA K, HOSHINO T, KNEBEL K, ET AL. BCNU-induced perturbations in the cell cycle of 9L rat brain tumor cells. Cancer Treat Rep 62:747–754, 1978.
5. HOSHINO T, ROSENBLUM ML, WILSON CB, ET AL. Chemotherapeutic implications of growth fraction and cell cycle time in glioblastomas. J Neurosurg 43:127–135, 1975.
6. LEVIN VA, CRAFTS D, NORMAN D, ET AL. Criteria for evaluating malignant brain tumor patients undergoing chemotherapy. J Neurosurg 47:329–335, 1977.
7. KAPLAN EL, and MEIER P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481, 1958.
8. GEHAN EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203–223, 1965.
9. LEVIN VA, HOFFMAN WF, PISCHER TL, ET AL. BCNU-5-fluorouracil combination therapy for recurrent malignant brain tumors. Cancer Treat Rep 62:2071–2076, 1978.

# Use of Right Atrial Catheter for Prolonged Iv Support in Cancer Patients [1,2]

**Ram K. Chillar,* Mark Farbstein, Owen B. Ellington, M. Ann Turner, John L. Fahey, Daniel U. Riihimaki, Cindy Smith, Georgina H. A. Fleming, Philip K. Zia, and Ernest Beutler [3]**

## SUMMARY

**A right atrial catheter has proven to be a well-tolerated technical advance for patients requiring prolonged vascular access. It is easily inserted and suited for ambulatory maintenance by the patient. Catheters are utilized for a wide spectrum of iv medications with an acceptably low complication rate. Most significantly, the infection rate is negligible, despite severely compromised hosts. Their use should be considered for any patient in whom problems with vascular access are anticipated. This report comprises our experience with the first 70 catheters in 66 patients in the Hematology/Oncology service.**

[Cancer Treat Rep 64:243–246, 1980]

Vascular access in patients with neoplastic diseases is frequently a serious technical problem. Frequent blood sampling, prolonged chemotherapy, antibiotics, blood components, and parenteral nutrition often exhaust accessible peripheral iv sites. Moreover, the repeated minor discomfort of venipunctures tends to sap the emotional reserve of the very ill patient. A right atrial catheter originally utilized for hyperalimentation in patients with chronic bowel disease (1) and subsequently used in several bone marrow transplantation programs, including those at the City of Hope National Medical Center (2), appears to be a satisfactory solution to these problems. This report summarizes our increasingly successful experience with the first 70 such catheters in 66 patients in the Hematology/Oncology service. It is often believed that when these catheters are used for parenteral alimentation their use for other purposes such as blood sampling, transfusion, and administration of antibiotics is prescribed. We have been able, however, to use the same catheter for all purposes requiring access to venous circulation and have therefore found it to be of great value in patient management.

## CATHETER INSERTION

A 90-cm elastic catheter with a Dacron cuff 30 cm from the proximal end is available in adult and pediatric sizes.[4]

The insertion technique was that described by Heimbach and Ivey (3). All of these catheters were inserted by a single team of two to three oncologic surgeons with no special training in vascular surgery. With experience, usually very little difficulty was encountered in the placement of the catheter. The proximal end is inserted under local anesthesia into the right cephalic vein whenever possible. The catheter tip is then advanced into the superior vena cava or right atrium under fluoroscopic control. Where fluoroscopy in the operating room is not available, intraoperative chest roentgenography should be performed to ascertain the correct position of the catheter tip. The distal end of the catheter is exteriorized through a subcutaneous tunnel, within which the Dacron cuff is placed (fig 1). It was necessary to insert the catheter into the left cephalic vein on several occasions, into the external jugular vein on two occasions, and into the internal jugular vein on one occasion. Catheter insertion failed in

[1]Received July 3, 1979; revised Oct 30, 1979; accepted Nov 26, 1979.

[2]Supported in part by the Martin B. Friedman Research Fund.

[3]Departments of Hematology (*R. K. Chillar, M. Farbstein, O. B. Ellington, M. A. Turner, J. L. Fahey,* and *E. Beutler*), General and Oncological Surgery (*D. U. Riihimaki*), Nursing (*C. Smith* and *G. H. A. Fleming*), and Pharmacy (*P. K. Zia*), City of Hope National Medical Center, Duarte, CA.

**Reprint requests to:* Dr. Ram K. Chillar, Division of Hematology-Oncology, Martin Luther King, Jr. General Hospital, 12021 S Wilmington Ave, Los Angeles, CA 90059.

[4]Parenteral Alimentation Catheter, Evergreen Medical Products, Inc, Medina, WA.

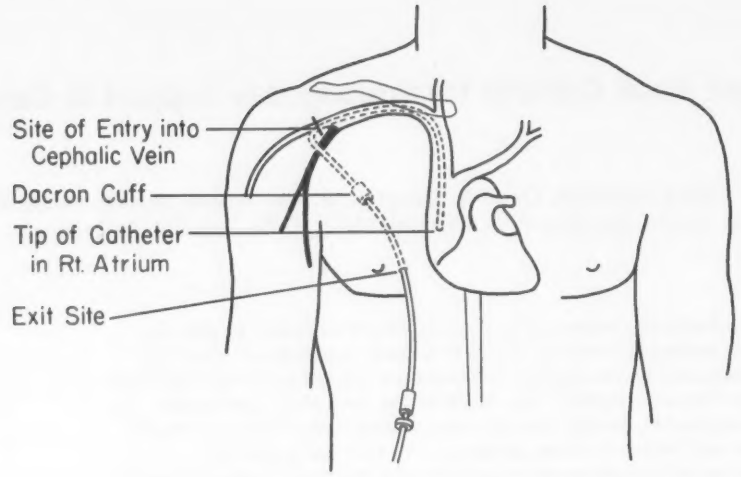



FIGURE 1.—Final position of catheter after placement is confirmed with fluoroscopy.

only one patient because a suitable vein could not be isolated. A vein large enough to accept an adult catheter was located in 62 of 70 cases. Pediatric size catheters were inserted in the other eight cases. During relapse, two patients needed a second catheter insertion in the opposite side at a later date. One patient had to undergo a third catheter insertion and the external jugular vein was used for this entry.

Catheters were inserted without complication despite severe granulocytopenia (wbcs $< 1000/mm^3$) in 22 patients and thrombocytopenia (platelets $< 50,000/mm^3$) in 25 patients. Prophylactic platelet transfusions were given to seven of the patients early in the course of this experience, but subsequently were not found to be necessary. When the need is desperate, it is now possible to perform the catheter insertions even in patients with severe thrombocytopenia using platelet support during and 24–72 hours after the operative period.

The anchoring skin suture was removed in approximately 10–14 days when fibrous tissue had organized around the subcutaneous cuff.

## CATHETER MAINTENANCE

Patients participated in catheter care while hospitalized and were capable of routine catheter maintenance by the time of discharge. The exit site skin was swabbed with alcohol and was cleaned daily with 3% hydrogen peroxide and dressed with Betadine (providone-iodine) ointment using a sterile technique. Catheters were flushed twice daily with 6 ml of saline containing 100 units of heparin/ml. Catheters were also flushed following blood sampling.

If resistance to free-flow or sluggish withdrawal of blood was noted, catheters were flushed with 12 ml of saline containing 100 units of heparin/ml or were flushed with saline containing 1000 units of heparin/liter and were infused with an infusion pump (McGaw Laboratories, Irvine, CA) at 75–100 ml/hour for 12 hours.

Catheters were capped with disposable sterile male Luer Locks (Berton-Dickinson and Co, Rutherford, NJ) after each entrance. The cost of the material used for routine maintenance is approximately $3.50/day.

## CATHETER REMOVAL

When catheters were no longer required, they were removed by surgical dissection of the tissue surrounding the Dacron cuff. On rare occasions it was possible to remove the catheter using firm traction only. Local pressure applied for 10 minutes prevents hematoma formation, and the subcutaneous tunnel is rapidly obliterated.

## RESULTS

During the past 2 years, 70 catheters have been inserted in 66 patients (table 1). Eleven patients with acute leukemia currently have catheters in place with durations of 8–395 days. The mean duration of catheter life was 85 days for 42 catheters inserted in patients with acute leukemia who either died (38 patients) or had their catheter removed (three patients). Catheters required in six pa-

TABLE 1.—Duration of 70 catheters inserted in 66 patients

| Diagnosis | Present patient status | No. of catheters inserted | Mean duration in days of catheter life (range) |
|---|---|---|---|
| Acute leukemia | Inpatient | 11 | 128 (8–395+) |
| | Outpatient | 42 | 85 (7–451) |
| Hodgkin's disease | Outpatient | 6 | 124 (57–240) |
| Testicular carcinoma | Inpatient | 1 | 60+* |
| | Outpatient | 6 | 76 (10–130) |
| Other | Outpatient | 4 | 101 (24–266) |

*Patient is still receiving chemotherapy through the catheter.

TABLE 2.—Reasons for removal of catheters

| | Catheters removed | |
|---|---|---|
| | No. | % |
| Blockage | 2 | 3.4 |
| Mechanical damage | 3 | 5.3 |
| Infection at exit site | 2 | 3.4 |
| Fever of undetermined origin | 1 | 1.7 |
| Other (leak-malposition) | 2 | 3.4 |
| Not needed | 18 | 31.1 |
| Death of patient | 30 | 51.7 |
| Total | 58 | 100.0 |

tients with Hodgkin's disease have all been removed after a mean duration of 124 days. Six patients with testicular carcinoma required catheters for a mean duration of 76 days, while another patient with testicular carcinoma is still receiving chemotherapy through the catheter after > 60 days. Four patients with nonmalignant disorders also needed these catheters for iv support for a mean duration of 101 days.

Catheters were used for routine blood sampling as well as for therapy with antineoplastic agents, antibiotics and other medications, fluids, and parenteral nutrition. Reasons for removal of 58 catheters are shown in table 2. Twelve patients still need these catheters for maintenance therapy.

## COMPLICATIONS

A variety of catheter-related complications have been encountered. Catheter clotting was the most frequent problem. Two catheters were removed because of complete blockage early in our experience. The use of a flexible guide wire to open clotted catheters was attempted in several cases and was uniformly unsuccessful.

Early in our experience, five catheters ruptured when excessive pressure was applied in attempts to re-establish patency. One catheter ruptured proximal to the subcutaneous tunnel and was therefore removed. One catheter was removed because it was inadvertently damaged during removal of the anchoring skin suture. Four catheters ruptured externally and were salvaged by suturing a 14-gauge blunt needle into the catheter lumen. More recently, streptokinase has been utilized in nine cases of complete catheter blockage. Aliquots of ≤ 6 ml of saline containing 250,000 units of streptokinase/6 ml were instilled and left in the catheter for 20 minutes. The solution was then entirely withdrawn and catheter flow was reevaluated. In all nine cases, catheter patency was reestablished. In one patient with fever and severe granulocytopenia, an episode of hypotension and bronchospasm was temporally related to streptokinase treatment for a completely occluded catheter. He subsequently required 3 days of vasosuppressor support. Although this may represent a complication of streptokinase, sepsis was also a clinical consideration. One patient developed pectoral hematoma requiring surgical evacuation in the immediate postoperative period. In one patient persistent fever necessitated the removal of the catheter. Culture of the discontinued catheter tip did not reveal any organisms and the patient continued to have a low-grade fever. We did not observe any increased incidence of infection in six patients whose catheters were also used for hyperalimentation in addition to chemotherapy, blood sampling, and antibiotics.

In one patient, a distal segment of catheter measuring 15 cm broke while the clotted catheter was being removed and lodged in the patient's right pulmonary artery. He died of leukemia relapse 170 days later and, at autopsy, the catheter segment was free of infection or deterioration.

## DISCUSSION

A right atrial catheter has proven to be a useful adjunct in the management of patients with a variety of malignant diseases. Vascular access, as well as repeated venipuncture, can become a major obstacle in the delivery of op-

timal care to such patients. When prolonged iv support or chemotherapy is necessary, a right atrial catheter should be considered.

In our experience, catheters can be inserted in almost all patients. No problems with hemorrhage were encountered despite the presence of severe thrombocytopenia in all except one patient who required evacuation of pectoral hematoma soon after the catheter placement. In no instance was the occurrence of septicemia clearly due to the presence of the catheter, although many patients had severe granulocytopenia and catheters were entered daily. Catheters were utilized for all modalities of iv therapy, including hyperalimentation, as well as for removal of blood samples. The low infection rate observed is undoubtedly related to meticulous nursing care and patient education. The long subcutaneous tunnel and Dacron cuff may also serve as a barrier to infection. Clotting within the catheter has been minimized by institution of a standard nursing protocol for routine catheter maintenance as well as for early intervention when sluggish flow develops. Those catheters which do become occluded may be salvaged by local treatment with streptokinase. It is possible that urokinase, a human enzyme, might be equally effective and less sensitizing. More significant complications, such as catheter wall rupture, have been eliminated as medical staff and patients have gained experience with these catheters. It is not surprising, therefore, that the last 15 catheters inserted in 1979 have been virtually free from all problems and streptokinase was not needed at all. Occasionally, a catheter tip becomes lodged against the vessel wall in such a way as to permit infusion of fluids but not withdrawal of blood. In this situation there is no intraluminal clot and clot lysis measures are not helpful.

## REFERENCES


1. Broviac JW, Cole JJ, and Scribner BH. A silicone rubber atrial catheter for prolonged parenteral alimentation. Surg Gynecol Obstet 136:602–606, 1973.
2. Blume KG, Bross KJ, Riihimaki DU, et al. The use of a right atrial catheter in bone marrow transplantation for acute leukemia. Exp Hematol 6:636–638, 1978.
3. Heimbach DM, and Ivey TD. Technique for placement of a permanent home hyperalimentation catheter. Surg Gynecol Obstet 143:634–636, 1976.

# Baker's Antifol in Combination With 5-Fluorouracil and Methyl-CCNU in the Treatment of Metastatic Colorectal Cancer: A Southwest Oncology Group Study (Protocol 7764)[1,2]

**Michael T. Shaw,[3,*] John D. Bonnet,[4] Henry Wilson,[5] and Lance K. Heilbrun[6,7]**

## SUMMARY

**Fifty-one previously untreated patients with metastatic colorectal cancer were treated with a combination of 5-fluorouracil, methyl-CCNU, and Baker's antifol. Six of 34 fully evaluable patients achieved a response. There were no apparent benefits obtained with this combination compared to treatment with 5-fluorouracil alone.**

[Cancer Treat Rep 64:247–250, 1980]

The treatment of metastatic colorectal cancer with chemotherapeutic agents has not been very satisfactory. 5-Fluorouracil (5-FU) produces a response rate of about 20% (1), which has varied from 8% to 85% (2). 5-FU and methyl-CCNU combinations with or without vincristine have produced response rates varying from 11% to 40% (3).

Baker's antifol (BAF) is a triazine antifol which is a potent inhibitor of dihydrofolate reductase. Phase I and II studies in man have demonstrated activity against adenocarcinoma of the colon. McCreary et al (4) noted an objective regression rate of 18% in 28 patients, 14 of whom had received previous chemotherapy. Padilla et al (5) described a partial response rate of 10% in 29 evaluable patients who had been previously treated with 5-FU alone or in combination with mitomycin C and methyl-CCNU.

Bone marrow toxicity has been shown to be mild after single doses, and it seemed logical to combine BAF with a 5-FU combination. A weekly schedule for the administration of BAF was based on a well-tolerated regimen established at Yale University, New Haven, CT (Bertino JR, personal communication). We chose 5-FU and methyl-CCNU to try to improve the response rate and survival of patients with metastatic colorectal cancer.

## MATERIALS AND METHODS

### Patients

A total of 51 patients were entered in this study between July 1977 and February 1979 when the study was closed. Fifty of these 51 patients were eligible according to the protocol criteria and also had measurable lesions. A total of 43 patients were either fully or partially evaluable for response, and they form the initial patient subsample upon which most analyses were performed. Followup for survival data is current as of April 30, 1979. Of the 43 fully and partially evaluable patients, 23 (53%) have died and the other 20 (47%) are still alive and under followup; no patients have been lost to followup. Their median followup time was 21 weeks. None of these patients had received previous chemotherapy, except one, who had received two doses of 5-FU. Untreated patients had an anticipated survival of at least 8 weeks, with adequate renal and hepatic functions as defined by a BUN of $< 20$ mg/100 ml, a serum creatinine of $< 1.5$ mg/100 ml, and a bilirubin of $< 1.5$ mg/100 ml. The initial total wbc count was $> 4.0 \times 10^9$/liter and the initial platelet count was $> 100 \times 10^9$/liter. Informed consent was obtained

[1]Received July 16, 1979; revised Oct 15, 1979; accepted Oct 25, 1979.

[2]Supported in part by Public Health Service grants CA-10187, CA-12014, CA-12213, and CA-21074 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Department of Medicine, University of New Mexico School of Medicine, Albuquerque.

[4]Scott and White Clinic and Hospital, Temple, TX.

[5]Ohio State University School of Medicine, Columbus.

[6]University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, and the Southwest Oncology Group Statistical Office, Houston.

[7]We wish to acknowledge the valuable technical assistance of Howard Conlon and Sue Wimberly.

*Reprint requests to: Michael T. Shaw, MD, Cancer Research Campaign Oncological Centre, Newcastle General Hospital, Westgate Rd, Newcastle upon Tyne, NE4 6BE, England.

from all patients entered in the study in accordance with institutional policies.

## Drug Administration

Methyl-CCNU was given at a dose of 60 mg/m² orally on Days 1 and 2 of each cycle, 5-FU was given at a dose of 400 mg/m² by iv push on Days 1–5 of each cycle, and BAF was given at a dose of 150 mg/m² by iv infusion over 45 minutes to 1 hour every week beginning on Day 1. The administration of methyl-CCNU and 5-FU was repeated every 6 weeks.

Patients who had received prior radiotherapy to > 25% of their bone marrow-bearing areas, who had tumor in their bone marrow, or who were > 65 years of age were considered poor risks and were given modified doses initially: 45 mg/m² of methyl-CCNU orally on Days 1 and 2 of each cycle, 300 mg/m² of 5-FU by iv push on Days 1–5 of each cycle, and 100 mg/m² of BAF by iv infusion over 45 minutes to 1 hour every week beginning on Day 1.

If the wbc count did not fall to $< 4.0 \times 10^9$/liter and the platelet count did not fall to $< 150 \times 10^9$/liter, doses were escalated accordingly. The maximum doses were 70 mg/m² of methyl-CCNU on Days 1 and 2 (dose was not increased more often than once every other course), 500 mg/m² of 5-FU on Days 1–5, and 150 mg/m² of BAF once weekly.

Intervals between courses were lengthened to allow for complete hematologic recovery before initiation of subsequent courses (wbc count $> 4.0 \times 10^9$/liter and platelet count $> 100 \times 10^9$/liter). Early leukopenia (before 21 days), a wbc count of $< 2.0 \times 10^9$/liter, required reduction of the dose of 5-FU by 25%. Late leukopenia (≥ 35 days), a wbc count of $< 2.0 \times 10^9$/liter, necessitated reduction of the dose of methyl-CCNU by 25%. Thrombocytopenia, a platelet count of $< 50 \times 10^9$/liter and $> 25 \times 10^9$/liter with no bleeding, required a reduction of the dose of methyl-CCNU by 25%. If the platelet count fell to $< 25 \times 10^9$/liter and/or significant bleeding secondary to thrombocytopenia occurred, the dose of methyl-CCNU was reduced by 50%. Stomatitis, judged clinically as severe, or diarrhea with more than five loose stools on any 1 day, required reduction of the dose of 5-FU by 25%. If the wbc count fell to $< 2.0 \times 10^9$/liter or the platelet count fell to $< 75 \times 10^9$/liter, BAF was omitted until the wbc count had reached $3.0 \times 10^9$/liter and the platelet count had reached $100 \times 10^9$/liter. If stomatitis, causing considerable discomfort to the patient, or dermatitis occurred, BAF was omitted until the lesions had disappeared.

An adequate trial consisted of 6 weeks of therapy, ie, one course of methyl-CCNU and 5-FU and six doses of BAF. Early death was defined as death within the first 4 weeks of therapy.

Therapy was continued until tumor progression, an increase of ≥ 25% in the sum of the products of the two maximal perpendicular diameters of the measured lesions or ≥ 30% in the sum of the measurements of the liver, was found. Complete remission (CR) was defined as the complete disappearance of all clinically recognizable evidence of tumor. Partial remission (PR) was defined as a decrease of measurement of ≥ 50% in the sum of the products of the two maximal perpendicular diameters of measured lesions or ≥ 30% in the sum of three or more liver measurements under the costal margin. Improvement was defined as a decrease of measurement of 25%–49% of the products of the two maximal perpendicular diameters of measured lesions or 15%–29% of the sum of three or more liver measurements under the costal margin.

## Statistical Methods

The method of survival curve estimation used was that of Kaplan and Meier (6). Survival comparisons between two groups of patients were made using the generalized Wilcoxon test due to Gehan (7). Two-sided tests of significance were performed except when comparing survival by response. In those cases, one-sided tests of significance were made because of the obvious a priori clinical assumptions of directional differences.

## RESULTS

Forty-three eligible patients entered in this protocol were evaluated; clinical details and therapeutic results are summarized in tables 1 and 2. Five patients died within 4 weeks of starting therapy, and four patients refused treatment after one course. Of the 34 patients having an adequate trial of therapy, there were two CRs and four PRs, with an overall response rate of 18%. Seven patients showed improvement. There were 21 treatment failures. The 95% confidence intervals shown in table 2

TABLE 1.—Clinical features of 43 eligible patients

| Clinical feature | No. of patients |
|---|---|
| Insufficient data | 8 |
| Partially + fully evaluable patients | 43 |
| Sex | |
| Men | 22 |
| Women | 21 |
| Age in yrs | |
| Median, 60 | |
| Range, 19–75 | |
| Metastatic sites | |
| Liver ± other sites | 27 |
| Lung, bone, intra-abdominal | 16 |
| Early deaths | 5 |
| Refused chemotherapy after 1 course | 4 |
| Fully evaluable | 34 |

TABLE 2.—Response data for fully evaluable patients

| | No. of patients |
|---|---|
| Total fully evaluable | 34 |
| Response | |
| CR | 2 |
| PR | 4 |
| Improvement | 7 |
| No response | 21 |
| CR + PR | 6 (18%)* |
| CR + PR + improvement | 13 (38%)† |

*95% confidence interval = 0.07, 0.35.
†95% confidence interval = 0.22, 0.56.

indicate that based on the available data, we are 95% certain that the true (unknown) CR + PR rate for this drug combination does not exceed 35%. A similar interpretation applies to the CR + PR + improvement rate.

One patient, a 64-year-old woman, was originally found to have peritoneal seeding of a colonic cancer which was the primary tumor being resected. Following three courses of 5-FU, methyl-CCNU, and BAF, she had intractable bleeding from a benign gastric ulcer and at reparative operation was found to be completely free of the peritoneal metastases, with no other evidence of tumor recurrence. She refused further treatment with cytotoxic drugs and thereafter remained well, with no outward evidence of recurrence 48 weeks following initiation of chemotherapy. Another patient, a 56-year-old woman whose primary lesion was rectal, was entered in the protocol after she developed pulmonary metastases which appeared on chest roentgenograms as multiple nodules. After two courses of 5-FU, methyl-CCNU, and BAF, several of the nodules disappeared and others were considerably smaller. After four more courses of therapy, all metastases had disappeared and remained so for 56 weeks following the start of treatment.

Survival data for the 43 fully or partially evaluable patients are shown in figure 1 (median, 32 weeks). Several other survival comparisons were also made. Survival of good-risk patients was not significantly different from that of poor-risk patients ($P = 0.357$). These two patient subgroups were previously defined. Four of the six responding patients have relapsed; the median response duration is 29 weeks. One patient is still in remission after 69 weeks. The survival of fully evaluable patients has been compared by response category in two ways. The survival of patients with CR + PR (median, 70 weeks) is significantly longer (one-sided, $P = 0.047$) than that achieved by the patients with improvement and no response (median, 34 weeks). Also, the survival of fully evaluable patients who have achieved improvement or better response (median, 73 weeks) is highly significantly longer (one-sided, $P = 0.001$) than that achieved by patients with no response (median, 26 weeks).

In most instances, toxicity was mild and certainly quite acceptable. Vomiting occurred in only one patient. Skin rashes of a pruritic-maculopapular type were seen in three patients. The rash in each case was most prominent on the trunk and arms and legs and appeared after the first infusion of BAF in two patients and after the second infusion in the other. In each instance BAF was stopped for 1 or 2 weeks until the rash had disappeared and was then re-started using full doses without reappearance of the skin lesions. Mild diarrhea which resolved spontaneously without modification of therapy was noted in three patients. Mild oral mucositis which resolved spontaneously appeared in three patients. Two cases of sepsis were recorded; one patient developed viral pneumonitis after the fifth course of therapy and died and another patient developed clostridial septicemia after the second course of chemotherapy and eventually died as a result of multiple abscesses.

Hematologic toxicity in most instances was not excessive and is summarized in table 3. After the third course of therapy, the platelet count nadirs tended to drop. One patient who had a platelet count nadir of 4000/mm$^3$ after the third course of treatment developed gastrointestinal bleeding and evidence of hemorrhage in her lungs. She was treated with rbc and platelet transfusions but died with evidence of massive multiorgan tumor infiltration.

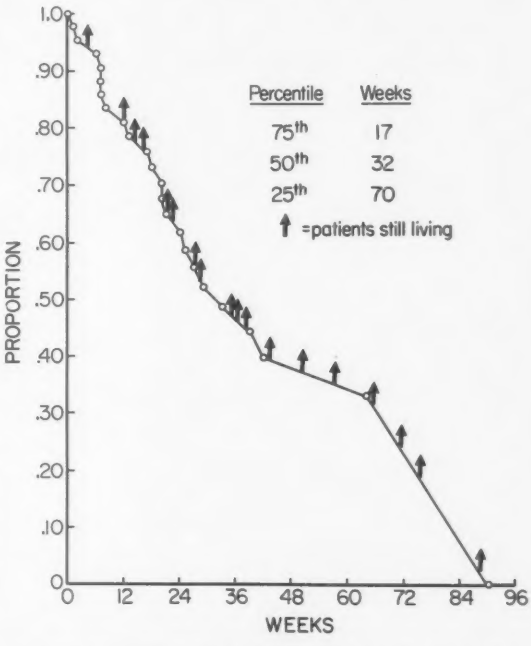



FIGURE 1.—Survival of fully and partially evaluable patients.

TABLE 3.—Hematologic toxicity

| No. of courses | No. of patients | Wbc count (cells × $10^9$/liter)* | Platelet count (cells × $10^9$/liter)* |
|---|---|---|---|
| 1 | 34 | 1.1–8.6 (3.5) | 10–374 (131) |
| 2 | 34 | 0.9–8.2 (3.5) | 26–407 (135) |
| 3 | 30 | 1.5–7.5 (3.2) | 4–496 (141) |
| 4 | 17 | 2.7–5.6 (3.8) | 70–230 (132) |
| 5 | 15 | 1.9–7.6 (3.4) | 74–159 (112) |
| 6 | 15 | 1.7–3.5 (2.7) | 60–164 (113) |
| 7 | 15 | 1.2–3.2 (2.6) | 25–193 (100) |
| 8 | 8 | 2.7–4.0 (3.2) | 57–160 (104) |

*Values given = range with mean in parentheses.

## DISCUSSION

The combination of methyl-CCNU, 5-FU, and BAF which was chosen for this treatment regimen of metastatic colorectal cancer seemed rational in view of the published advantage of methyl-CCNU-5-FU combinations over 5-FU alone in the response rate of advanced disease. Furthermore, BAF has been shown to have activity in the disease, even when patients had been previously treated with other chemotherapeutic agents. Nevertheless, in our study, the response rate of 18% was no better than that achieved following the administration of 5-FU alone. Although studies (8,9) have demonstrated the superiority of the combination of methyl-CCNU and 5-FU over 5-FU alone Moertel (3) has recently rejected this alleged advantage of the two-drug combination. Our three-drug combination does not promise to improve the existing abysmally low response rate of metastatic colorectal cancer. However, we were able to administer the drugs without undue toxicity in most cases and to reduce the discomfort caused by vomiting by giving methyl-CCNU on 2 successive days instead of in one dose. At the present time, the search for an active chemotherapeutic agent for cancer of the colon and rectum remains an elusive one.

## REFERENCES


1. CARTER SK. Guest editorial: Large-bowel cancer—the current status of treatment. JNCI 56:3–10, 1976.
2. MOERTEL CG, and REITEMEIER RJ. Advanced gastrointestinal cancer. *In* Clinical Management and Therapy. New York, NY, Harper and Row, 1969.
3. MOERTEL CG. Current concepts in cancer: chemotherapy of gastrointestinal cancer. N Engl J Med 299:1049–1052, 1978.
4. MCCREARY RH, MOERTEL CG, SCHUTT AJ, ET AL. A phase II study of triazinate (NSC 139105) in advanced colorectal carcinoma. Cancer 40:9–13, 1977.
5. PADILLA F, CORREA J, BUROKER T, ET AL. Phase II study of Baker's antifol in advanced colorectal cancer. Cancer Treat Rep 62:553–554, 1978.
6. KAPLAN E, and MEIER P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481, 1958.
7. GEHAN EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203–223, 1965.
8. BAKER LH, VAITKEVICIUS VK, and GEHAN E. Randomized prospective trial comparing 5-fluorouracil (NSC-19893) to 5-fluorouracil and methyl-CCNU (NSC-95441) in advanced gastrointestinal cancer. Cancer Treat Rep 60:733–737, 1976.
9. MOERTEL CG, SCHUTT AJ, HAHN RG, ET AL. Brief communication: Therapy of advanced colorectal cancer with a combination of 5-fluorouracil, methyl-1,3-*cis*(2-chlorethyl)-1-nitrosourea, and vincristine. JNCI 54:69–71, 1975.

# Clinical Toxic Effects of 2,3-Dihydro-1*H*-imidazo[1,2-*b*]pyrazole (IMPY) With Relevant Pharmacokinetic Parameters [1,2]

James A. Neidhart,* Alfred E. Staubus, Donn Young, Gerald W. King, and Louis Malspeis [3,4]

## SUMMARY

**2,3-Dihydro-1*H*-imidazo[1,2-*b*]pyrazole (IMPY) is an inhibitor of ribonucleotide reductase and of DNA synthesis selected for clinical trials because of its activity against L1210 leukemia variants resistant to other inhibitors of this enzyme. A phase I trial designed to allow in-depth pharmacologic evaluation has recently been completed and the clinical results and preliminary pharmacokinetic data are reported here. Each patient received IMPY by three different schedules. A single iv bolus, intermittent 5-day bolus, and 5-day continuous infusion were given at 3-week intervals. The major dose-limiting toxic effects were vomiting, rbc hemolysis, confusion, and somnolence. All toxic effects seemed to be dose- and schedule-dependent and were readily reversible. IMPY enters the cerebrospinal fluid and is highly concentrated in gastric secretions. Clearance of IMPY is impaired in the presence of hepatic insufficiency. Eighteen of 26 patients entered are evaluable for response, including one patient with colon cancer with minimal response and three patients with stable disease.**

[Cancer Treat Rep 64:251–255, 1980]

2,3-Dihydro-1*H*-imidazo[1,2-*b*]pyrazole (IMPY) produces rapid and readily reversible inhibition of DNA synthesis in vitro but has little effect on RNA or protein synthesis (1). Observations that the drug synchronizes HeLa cells at entry in the S phase (2) and is active against L1210 variants resistant to other inhibitors of ribonucleotide reductase[5] led to its introduction into clinical trials. This trial was designed to allow determination of the pharmacokinetics of IMPY coincident with evaluation of clinical toxicity. The toxic effects of IMPY are dose-related, predictable, and differ from those seen with most other anticancer agents. The pharmacologic observations reported here may help explain the toxic effects and allow better design of any future trials of IMPY.

## MATERIALS AND METHODS

Twenty-six patients with advanced malignant disease of a variety of histologic types were entered in the study. Patient characteristics are shown in table 1. All but one of these patients had progressive disease after radiation therapy or chemotherapy and all gave informed consent. Most patients had multiple-organ involvement by metastases, with a mean of 2.5 tumor sites per patient. Eleven patients had abnormal liver scans and nine had mild elevations of liver enzymes at the time of entry in the study, but none had elevated bilirubin. All patients had normal renal functions. Initially, one patient had compromised marrow reserve from metastatic prostatic cancer. Blood cell counts and liver enzymes were obtained daily during drug therapy and weekly between treatments. In selected patients measurements of serum hap-

[1]Received July 17, 1979; revised Oct 15, 1979; accepted Oct 29, 1979.

[2]Supported by contract N01-CM-53828 (Division of Cancer Treatment) and Public Health Service grant CA-16048 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Division of Hematology and Oncology, College of Medicine (*J. A. Neidhart* and *G. W. King*), Division of Pharmaceutics and Pharmaceutical Chemistry, College of Pharmacy (*A. E. Staubus* and *L. Malspeis*), and the Comprehensive Cancer Center (*D. Young*), The Ohio State University Hospitals, Columbus.

[4]We wish to thank the nurses of the Interdisciplinary Oncology Unit for the patient care and quality control that made this study possible, and Joan Loomis for her secretarial assistance.

[5]Clinical brochure: pyrazolo-imidazole (IMPY) (NSC 51143). Prepared by the Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD, Aug 1977, 69 pp.

**Reprint requests to*: James A. Neidhart, MD, Division of Hematology and Oncology, The Ohio State University Hospitals, 410 W 10th Ave, Rm N-1022, Columbus, OH 43210.

TABLE 1.—Patient characteristics

| Characteristic | No. of patients |
|---|---|
| Total | 26 |
| Male/female | 14/12 |
| Mean age in yrs (range) | 52 (19–74) |
| Prior chemotherapy | 25 |
| Prior radiation therapy | 13 |
| Evaluable for toxicity | 26 |
| Evaluable for response | 18 |

toglobin, urine hemosiderin, rbc osmotic fragility, methemoglobin levels, Heinz bodies, and adenosine triphosphate (ATP) levels were obtained near the end of the study. Tumor measurements were done monthly.

The schedules for IMPY administration were designed to determine and test pharmacokinetic parameters. The initial treatment was by single rapid iv bolus. Three weeks later patients received five daily rapid iv infusions of IMPY, and after another 3-week interval they received a 5-day continuous infusion. Starting doses and escalations are shown in table 2. Doses for continuous infusion were calculated from pharmacokinetic data (clearance) obtained from the first two schedules with the goal of maintaining a steady-state plasma concentration of IMPY equal to the plasma level seen 3 hours after a single bolus. If IMPY therapy was warranted beyond the initial three courses, the schedule used was selected by evaluating the clinical and pharmacokinetic data available for that patient. All patients received at least one treatment using this single bolus schedule, with a total of 36 courses given by that schedule. Seventeen patients received one course using the 5-day intermittent schedule and 14 patients received a third course of therapy using a continuous infusion.

IMPY levels were measured using an assay developed by L. Malspeis. An internal standard of 7-bromo-IMPY was added to each sample followed by extraction into methylene chloride. The extracted IMPY was then derivatized by adding heptafluorobutyric anhydride. The samples were evaporated to dryness, reconstituted with benzene, and washed with saturated sodium borate solution. Analyses were carried out using a Hewlett-Packard 5714A gas chromatograph equipped with a $^{63}$Ni electron-capture (pulsed, variable frequency type) detector. Chromatographic columns were 183 cm × 2-mm internal diameter coiled borosilicate glass packed with 3% OV-210 on 100–120-mesh Gas-Chrom Q (Applied Science Laboratories). This assay measures only the parent compound with a sensitivity of approximately 1 ng/ml in biologic specimens.

## RESULTS

The dose-limiting toxic effects were vomiting, abrupt drop in hemoglobin, and changes in mental status. Hematemesis is listed as a major toxic effect since it may have contributed to the fall in hemoglobin, although it usually represented only guaiac-positive emesis or streaking with blood and in no case required specific therapy. Table 2 shows the correlations of these toxic effects to dose, schedule, and each other. Nausea and vomiting were noted first and were the least-tolerated side effects. Confusion and obtundation were the most common central nervous system changes, although one patient developed psychosis. These symptoms were seen only at higher doses in patients with severe vomiting and a decrease in hemoglobin. The central nervous system changes always resulted in cessation of therapy, and slowly resolved over a period of 24–72 hours. There was a marked schedule dependency of these toxic effects. Total doses of 4000 and 9000 mg/m$^2$/course were easily tolerated with the 5-day intermittent and continuous-infusion schedules respectively, while 2700 mg/m$^2$ given as a single bolus produced dose-limiting toxicity.

Anemia, as evidenced by a fall in hemoglobin of > 3 g/dl, occurred in 23 cases and was accompanied by reticulocytosis in 17 cases. The fall in hemoglobin was abrupt, with onset usually within 24 hours of a single bolus or within 72 hours of 5-day therapy, with a mean time to nadir of 5.8 days. In nine cases the hemoglobin fell by > 5 g (maximum fall, 6.5 g) before transfusions were initiated. In no case did the degree of gastrointestinal blood loss seem proportionate to the hemoglobin drop, and during the latter part of the study we attempted to determine if hemolysis was a factor. Peripheral blood smears showed occasional fragmentation of rbcs but were otherwise normal. Two patients with severe decreases in hemoglobin had hemosiderinuria and one patient had hemoglobinuria. LDH was not serially measured in most patients, but was acutely and transiently elevated to fourfold and eightfold normal values in the two patients with hemoglobin decreases who received a 4-g/m$^2$ bolus. Serum haptoglobin values fell in one prospectively studied patient who developed anemia. Bone marrow examination in three patients showed normoblastic erythroid hyperplasia shortly after anemia developed. We were unable to demonstrate Heinz bodies, increased osmotic fragility of rbcs, or depletion of glucose-6-phosphate dehydrogenase. Rbc ATP levels (normal, 4.0–6.8 μM ATP/g of hemoglobin) were 3.4 and 3.8 μM in two patients shortly after a fall in hemoglobin. ATP values dropped from 5.2 to 2.3 μM and from 4.3 to 2.6 μM in two patients followed serially during treatment. Methemoglobin levels were prospectively obtained in one patient who did not develop hemolysis and whose hemoglobin rose from 0 to 3.2%, suggesting a mild and nonlethal oxidative stress on her rbcs.

Other toxic effects included oliguric renal failure in four patients, with the highest creatinine value being 3.5 mg/dl. This occurred only at higher doses in patients with

TABLE 2.—Major toxic effects of IMPY by dose and schedule

| Schedule | Dose (mg/m²/day) | No. of patients/ No. of courses | No. of patients with— | | | |
|---|---|---|---|---|---|---|
| | | | Vomiting* | Hematemesis | Hemolysis* | Central nervous system changes |
| Single bolus | | | | | | |
| | 150 | 2/2 | 0 | 0 | 0 | 0 |
| | 500 | 5/6 | 0 | 0 | 0 | 0 |
| | 800 | 9/12 | 0 | 0 | 0 | 0 |
| | 1200 | 1/1 | 0 | 0 | 0 | 0 |
| | 1800 | 4/4 | 3 (2) | 0 | 0 | 0 |
| | 2700 | 9/9 | 2 (2) | 2 | 2 (0) | 0 |
| | 4000 | 2/2 | 2 (2) | 2 | 2 (2) | 2 |
| 5-day intermittent bolus | | | | | | |
| | 150 | 1/1 | 0 | 0 | 0 | 0 |
| | 500 | 2/2 | 1 (0) | 0 | 0 | 0 |
| | 800 | 4/4 | 3 (0) | 0 | 1 (0) | 0 |
| | 1200 | 6/6 | 2 (1) | 1 | 2 (0) | 0 |
| | 1800 | 4/4 | 4 (4) | 3 | 4 (2) | 1 |
| 5-day continuous infusion | | | | | | |
| | 1800 | 2/2 | 1 (0) | 1 | 1 (0) | 1 |
| | 2300 | 10/10 | 9 (6) | 8 | 9 (4) | 2 |
| | 3000 | 2/2 | 2 (2) | 2 | 2 (1) | 2 |

*Nos. in parentheses = No. of patients with *severe* toxic effects: vomiting of ≥ 6 episodes/day and hemolysis of ≥ 5 g/dl decrease in hemoglobin.

severe hemolysis. Patients were treated with forced diuresis, and renal function returned to normal within several days. Two of these patients developed jaundice (highest bilirubin level, 14.1 mg/dl) and one patient had a threefold elevation of liver alkaline phosphatase which suggested an obstructive component. Elevations of SGOT and SGPT without a rise in alkaline phosphatase or bilirubin were seen in four additional patients coincident with hemolysis. All of the above toxic effects were acute, occurring within 5 days of initiation of therapy and correcting within 4 days of cessation of therapy. Malaise and fatigue were major problems at higher doses for several days after therapy. Miscellaneous toxic effects included constipation (two cases), diarrhea (two cases), hypotension (five cases), prolonged bleeding time (one case), and hiccoughs (three cases). No evidence of bone marrow suppression was seen.

Comparative plasma concentration-time curves of IMPY in a patient with hepatic metastasis who developed obstructive jaundice just before her third course of treatment are shown in figure 1. Although the volume of distribution did not change substantially, the rapid disposition phase half-life increased from 3.8 to 14.6 minutes. The terminal biologic half-life increased from 1.6 to 5.0 hours for the third course. Total-body clearance similarly decreased from 130 to 46 ml/minute/m² after development of jaundice. The pharmacokinetics of the first and second courses was almost identical suggesting that fairly severe liver dysfunction was necessary to alter clearance. This degree of change cannot be attributed solely to the 500-mg/m² increase in dose for the third course. The terminal biologic half-life in eight other patients treated with either 500 or 650 mg/m² ranged from 1.7 to 3.9 hours, with a mean of 3.2 hours. Total-body clearance remained > 61 ml/minute/m², with a mean value of 113.5 ml/minute/m².

Serial gastric fluid samples were available in another patient and showed substantially higher concentrations of IMPY in gastric secretions than in plasma. Sixty minutes following an iv bolus of 500 mg/m², plasma levels were 16.5 μg/ml and gastric fluid levels were 1054.3 μg/ml, representing a 64-fold concentration gradient. Five hours later plasma levels were 4.2 μg/ml and gastric fluid levels were 150 μg/ml. One patient who developed confusion while receiving IMPY by continuous infusion at 1800 mg/m²/day required a diagnostic lumbar puncture. Cerebrospinal fluid levels of IMPY were 23.1 μg/ml while concurrent steady-state plasma levels were 30.8 μg/ml. In this patient, cerebrospinal fluid IMPY levels were 75% of serum levels.

In a patient with colon cancer, an abdominal mass measured initially at 22 × 9 cm on physical examination decreased to 8 × 8 cm and may represent a minimal response. This patient received 1800 mg/m² × 5 days and refused further therapy due to vomiting. One patient with renal cell carcinoma required radiation therapy to

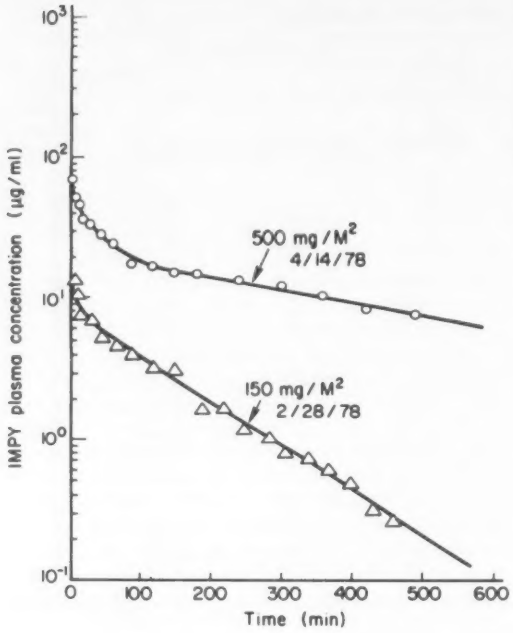



FIGURE 1.—Comparative plasma concentration-time curves in a patient who developed jaundice prior to the 3rd course of IMPY. $t_{1/2}\beta$ for the 1st course was 1.59 hrs (2/28) increasing to 5.0 hrs with the 3rd course (4/14).

a mass occupying approximately one half of the abdomen and extending into the thigh. She had been unresponsive to three previous courses of IMPY but received the radiation therapy concurrently with a fourth course with a 75% reduction in tumor following 1200 rads of radiation. This is an unusually good response to radiation therapy in renal cell carcinoma and may be a hint of synergism. Nonresponders included four patients with colon cancer, two with renal cell cancer, two with melanoma, and one each with eight other tumor types.

## DISCUSSION

Preclinical testing showed that IMPY had moderate activity against L1210 leukemia, including variants resistant to DTIC, guanazole, hydroxyurea, and cytosine arabinoside. IMPY demonstrated marginal activity against murine adenocarcinoma, colon cancer, melanoma, and osteogenic sarcoma. The mechanism of this activity was thought to be related to inhibition of ribonucleotide reductase (1), although the antitumor effect of IMPY is not schedule-dependent as is that of other inhibitors of this enzyme (guanazole and hydroxyurea) and IMPY is active against cell lines resistant to these other agents (*see* footnote 5). The formation of an active metabolite or a second biologic activity of IMPY might explain these findings. Brockman et al (3) shed some light on this issue by demonstrating that inhibition of DNA synthesis by IMPY is prolonged to 18–20 hours after in vivo administration (3) contrasted with the rapidly reversible in vitro effect seen in drug-wash-out experiments (2). This finding suggested that a metabolite might contribute to inhibition of DNA synthesis following IMPY treatment. We agree and have recently demonstrated cell synchronization and inhibition of DNA synthesis by the β-ribosyl metabolite of IMPY (4). This becomes important since the acute toxic effects of IMPY are schedule-dependent, suggesting that a dose schedule producing maximal S-phase block by metabolite with minimal toxicity by the parent compound (eg, a 5-day intermittent schedule) might be feasible.

It is clear, however, that IMPY has important and unique biologic activities aside from inhibition of DNA synthesis. This agent seems capable of inflicting lethal damage on mature non-nucleated rbcs and may inflict similar injury on other normal or malignant cells. Although the nature of this injury is not clear, we have evidence of generation of hydrogen peroxide in rbcs in vitro (5) and evidence of oxidant stress by methemoglobin formation in vivo. Severe oxidant stress associated with methemoglobinemia might produce hemolysis (6), although a causative relationship of oxidative stress to the abrupt anemia in our patient is still somewhat speculative. The significance of the fall in rbc ATP levels in patients receiving IMPY is not clear to us. If the unique and potent biologic effect of radical formation by IMPY proves to be exploitable for anticancer activity, the observation must be classified as serendipity since it is difficult to relate to the original reasons for developing this agent.

The onset of serious toxicity with dose escalation of IMPY is fairly sharp. With a single injection every 3 weeks, doses of between 1200 and 1800 mg/m² should not produce severe toxicity. Many patients, however, will not be able to tolerate doses of 2700 mg/m² and escalation will be difficult. The total dose per course can be increased fivefold (eg, 1200 mg/m²/day for 5 days) with little increase in toxicity. Since IMPY has a relatively long terminal half-life of approximately 7 hours at clinically useful doses (7), and since a metabolite may be an inhibitor of DNA synthesis, there seems to be little rationale for use of a continuous infusion. The gastrointestinal toxicity may well be related to the high concentration of IMPY in gastric secretions and peak serum IMPY levels, and therefore, schedule-dependent. It is unlikely that objective toxic effects other than mild hemolysis will be a problem unless escalation to doses producing severe vomiting is attempted. Patients with hepatic dysfunction should probably have a reduction in initial drug dose.

Nontoxic dose levels should still be effective in terms

of inhibition of DNA synthesis for the reasons mentioned above. We have found that IMPY produces hydrogen peroxide in rbcs at plasma concentrations attainable without marked gastrointestinal toxic effects in most patients (5). In addition, one patient in the present study developed methemoglobinemia as evidence of oxidative stress at nontoxic doses of IMPY. It is quite possible that this agent can produce useful antitumor effects at nontoxic doses. The lack of bone marrow suppression and potent oxidative stress induced by this drug raise the possibility of synergism with other chemotherapeutic agents and radiation therapy.

## REFERENCES


1. Ennis HL, Möller L, Wang JJ, et al. 2,3-Dihydro-1*H*-imidazole(1,2-b)pyrazolo: a new inhibitor of deoxyribonucleic acid synthesis. Biochem Pharmacol 20:2639–2646, 1971.
2. Beer CT, Kajiwara K, and Mueller GC. Synchronization of HeLa cells with 2,3-dihydro-1*H*-imidazole(1,2-*b*)pyrazole. Biochem Pharmacol 23:1115–1122, 1974.
3. Brockman RW, Shaddix SC, Carpenter JW, et al. Studies on the mechanism of action of 2,3-dihydro-1*H*-pyrazolo[2,3-*a*]imidazole (NSC-51143) and evidence of its metabolism in vivo. Proc Am Assoc Cancer Res and ASCO 19:21, 1978.
4. Neidhart J, and Sagone A. IMPY and its β-ribosyl metabolite show differential in vitro effects. Proc Am Assoc Cancer Res and ASCO 20:177, 1979.
5. Sagone AL, Malspeis L, and Neidhart JA. Generation of reactive oxygen compounds by new antitumor agent, IMPY (NSC 51143). Proc Am Fed Clin Res 27:391A, 1979.
6. Chan TK, Mak LW, and Ng RP. Methemoglobinemia, Heinz bodies, and acute massive intravascular hemolysis in lysol poisoning. Blood 38:739–744, 1971.
7. Staubus AE, Kerr TA, Balcerzak SP, et al. Pharmacokinetics of pyrazolo-imidazole (IMPY, NSC 51143) in humans. Proc Am Assoc Cancer Res and ASCO 20:366, 1979.

# Neurotoxicity of Adriamycin in Rats: A Low-Dose Effect[1,2]

**Bernard S. Jortner and Eun-Sook Cho**[3,4,*]

## SUMMARY

**A neurotoxic effect of Adriamycin in rats has been noted following a single iv dose level of 5 mg/kg of body weight (35 mg/m$^2$ of body surface area). The changes which were studied in the dorsal root ganglia were qualitatively similar to those previously reported at a higher dose level. The early changes involved neuronal nuclei and included focal regions of clearing and altered staining of chromatin. Necrosis of many affected neurons followed, but the evolution of this lesion was slower than that seen in previous studies using the higher dose of the drug.**

[Cancer Treat Rep 64:257–261, 1980]

Adriamycin has become an important cancer chemotherapeutic agent and is useful in treating a variety of neoplasms (1,2). A neurotoxic effect of this drug in rats given a single iv dose of 10 mg/kg of body weight was clinically noted by Philips et al (3), who reported progressive asynergia of the hind limbs. Cho (4) described the morphologic substrate of this abnormality, indicating that neuronal necrosis in peripheral ganglia was the primary lesion. Recent studies using the 10-mg/kg dose indicated that the initial morphologic changes in affected neurons took place in the nucleus, and consisted of focal clearing due to local reductions in chromatin (5). These were thought to represent a morphologic manifestation of the direct effect of the drug on chromosomal DNA (5), a known site of action of this compound (6). This paper reports the neurotoxic effect of 5 mg/kg of body weight (35 mg/m$^2$ of body surface area) of Adriamycin in rats, a dose compatible with prolonged survival of the animals (3), and suggests a potential for similar toxic effects in man.

## MATERIAL AND METHODS

Male Sprague-Dawley rats (Marland Farms) weighing approximately 250 g were anesthetized with ip pentobarbital sodium. The femoral vein was exposed surgically and Adriamycin in a physiologic saline diluent was administered iv at a dose of 5 mg/kg of body weight, after which the incision was closed with stainless-steel wound clips. Controls received equivalent amounts of the diluent. Clinical observations for neurologic abnormalities were performed during the course of the study.

Animals were randomly selected for perfusion fixation at 2 (one control and two experimental animals), 15 (one control and three experimental), 28 (two experimental), and 49 (one control and one experimental) days after drug administration. Following induction of anesthesia as noted above, a perfusate of 50 ml of 4% paraformaldehyde, followed by 500 ml of 2.5% glutaraldehyde in Sorenson's phosphate buffer (pH 7.35), was administered via a cannula placed in the root of the aorta. The dorsal root ganglia and spinal nerve roots were dissected out, placed in the glutaraldehyde solution for 3–10 hours, rinsed in the buffer, post-fixed in 1% osmium tetroxide, and embedded in Epon. One-μm thick sections were cut, stained with toluidine blue, and examined by light microscopy.

## RESULTS

No overt neurologic signs were noted in animals killed at 2 and 15 days after drug administration. Equivocal unsteadiness was seen prior to euthanasia in the two

[1]Received July 24, 1979; revised Nov 26, 1979; accepted Nov 28, 1979.

[2]Supported by grant IN92G from the American Cancer Society.

[3]Department of Pathology, College of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark.

[4]The assistance of Benedict Albanese is gratefully acknowledged.

*_Reprint requests to:_ Eun-Sook Cho, MD, Department of Pathology, College of Medicine and Dentistry of New Jersey, New Jersey Medical School, 100 Bergen St, Newark, NJ 07103.

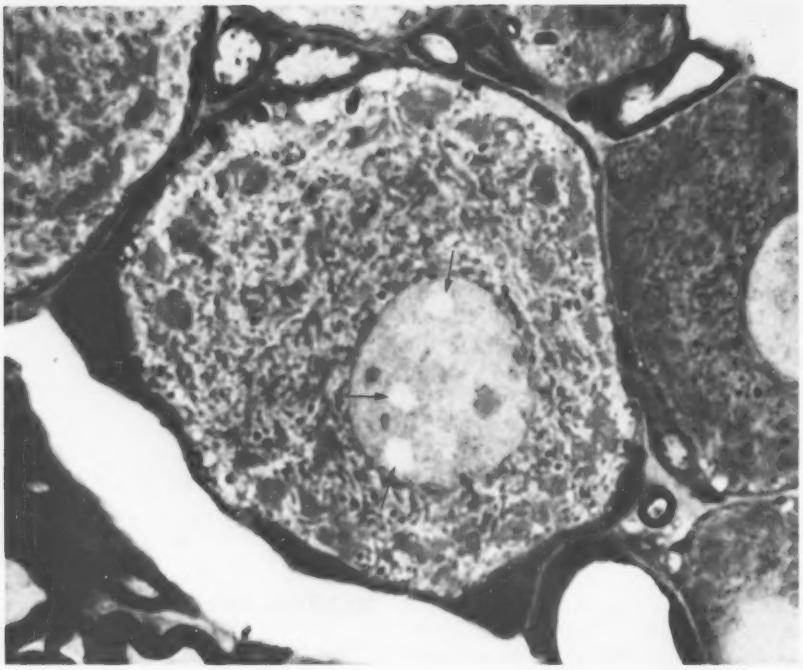


FIGURE 1.—Dorsal root ganglion neuron 2 days after iv administration of 5 mg/kg of Adriamycin. Arrows = focal regions of nuclear clearing (toluidine blue stain, × 1100).

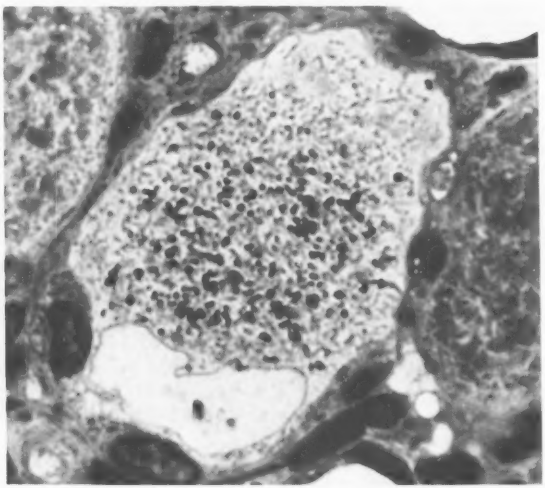


FIGURE 2.—Dorsal root ganglion neuron 15 days after administration of Adriamycin. The pale staining nucleus is peripheral in the cell soma and many darkly stained cytoplasmic dense bodies are evident (toluidine blue stain, × 1050).

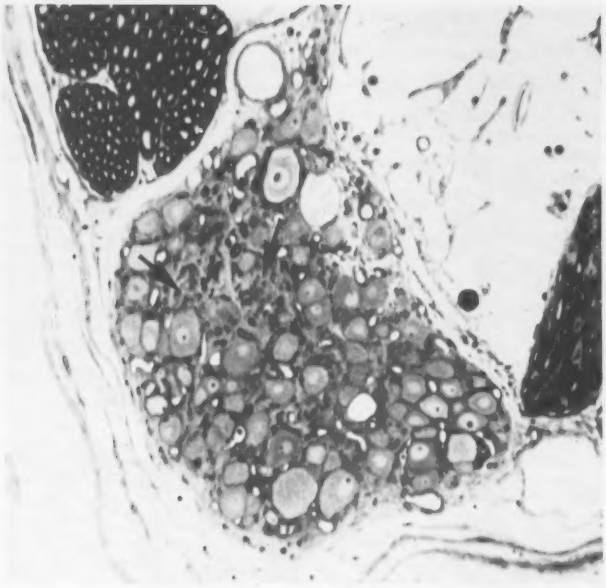


FIGURE 3.—Section of a dorsal root ganglion 28 days after drug administration. Arrows = regions of reactive nodules of Nageotte (toluidine blue stain, × 137).

experimental animals killed on Day 28. The single experimental animal perfused on Day 49 showed ataxia, most prominent in the hind limbs, and diminished response to pin-prick and touching of vibrissae.

A spectrum of light microscopic changes was noted in the dorsal root ganglia. While all ganglia and animals were not equally involved, a general progression of the lesions could be detected. In animals examined 2 days after the administration of Adriamycin, lesions were restricted to nuclei of neurons and consisted of focal round zones of clearing (fig 1) that were observed in about 70%–80% of the cells. Stainable chromatin aggregates appeared to be reduced in size, giving nuclei a "granular" appearance. Little significant cytoplasmic abnormality was observed.

By 15 days after drug administration, neuronal changes progressing to cell necrosis were well-developed in dorsal root ganglia. Nuclear alterations were prominent and included pallor and granularity of staining, zones of clearing, peripheral location in the cell soma, and karyolysis (fig 2). Cytoplasmic changes were observed in some cells with affected nuclei. These included an increase in the number of dense staining granules (fig 2), often perinuclear in location, and pale staining whorled regions of cytoplasm, probably representing collections of neurofilaments (4). In addition, some cells contained many large cytoplasmic vacuoles. Groups of proliferated satellite cells formed around and replaced necrotic neurons. Such cellular aggregates are termed nodules of Nageotte. From these changes it appeared that 5%–10% of neurons in affected ganglia had undergone necrosis.

In dorsal root ganglia from rats studied 28 days after drug administration, nuclear and cytoplasmic changes were qualitatively similar to those noted at 15 days, although the nodules of Nageotte were much more prominent (fig 3). About 20%–30% of neurons appeared to have undergone necrosis. Axonal swelling and fragmentation of myelinated nerve fibers into "digestion chambers" resembling those of wallerian degeneration were prominent in dorsal spinal nerve roots (fig 4).

By 49 days, at which time the experimental animal studied showed ataxia, smaller numbers of neurons had nuclear and cytoplasmic changes noted above. Instead, nodules of Nageotte and wallerian-like degeneration were more prominent. About 40% of the neurons were lost.

At all stages studied it was estimated that the proportion of neurons involved was fairly constant. However, the number of cells showing only nuclear changes diminished as the study progressed and the number of cells undergoing necrosis increased.

## DISCUSSION

We have demonstrated that qualitatively similar dorsal root ganglioneuronopathy in rats can be induced with 5

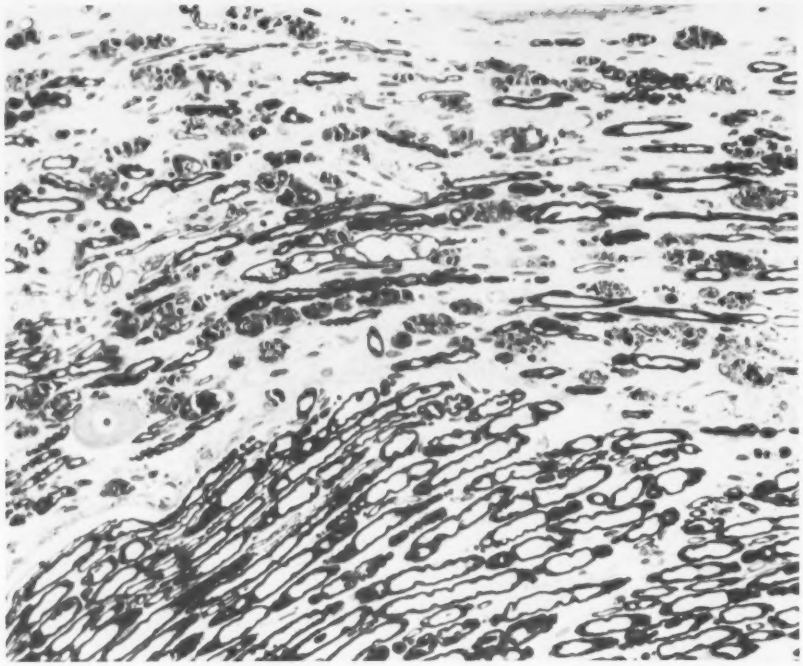


FIGURE 4.—Severe wallerian-like degeneration of fibers of the dorsal spinal nerve root in a rat 28 days after drug administration is present in the upper portion of the fig, in contrast to the well-preserved ventral root in the lower portion (toluidine blue stain, × 172).

mg/kg of Adriamycin, one half the dose used previously (5). This 5-mg/kg dose is equivalent to 35 mg/m$^2$ of body surface area (7), and is considerably lower than the 60–75-mg/m$^2$ individual doses administered triweekly in the recommended therapeutic schedules for man (1,2). In the present study, neuronal nuclear changes such as focal clearing were noted in the rats 2 days after drug administration, the earliest interval studied. These changes were undoubtedly present before that time, and could have had an onset similar to that seen in the 10-mg/kg dose study where such lesions were seen at 3 hours (5). Subsequent morphologic changes in dorsal root ganglia neurons evolved at a relatively slow rate in the present low-dose study. Fifteen days after administration a loss of about 5%–10% of neurons was noted. The neuronal loss rose to about 20%–30% at 28 days and 40% at 49 days, a point at which definite clinical signs were present in a single animal. In contrast, Cho et al (5) found necrosis of about 40% of the neurons in dorsal root ganglia of rats 15 days after administration of 10 mg/kg of Adriamycin, along with an earlier (11–12 days) onset of ataxia.

The selective neurotoxic effect of Adriamycin on neurons of peripheral ganglia of the rat has been noted (4,5) and is thought to be related to the permeability of blood vessels in such regions of the nervous system (8,9). After leaving the vessels the drug enters neurons and gains access to their nuclei.[5] Based upon evidence from other experimental systems, it likely intercalates between adjacent base pairs of DNA, inducing structural and functional changes such as alteration of the helical structure and breaks in the DNA molecule, as well as inhibition of DNA and RNA synthesis (6,10). The early neuronal nuclear changes noted in the present study may be a morphologic representation of some of these events. Since replication and repair synthesis of nuclear DNA in mature neurons is apparently absent (11), Cho et al (5) have postulated that the Adriamycin-induced damage eventually leads to cell necrosis. The presence of more slowly evolving neuronal necrosis in animals given 5 mg/kg as compared to those receiving 10 mg/kg (5) may indicate that the extent of DNA damage in such cells is propor-

[5]Mendell JR, and Sahenk Z. Abnormal neuronal perikaryon processing of $^3H$-leucine in Adriamycin treated rats. Presented at the American Academy of Neurology Meeting, Los Angeles, CA, 1978.

tional to drug dose and determines the rate of progression of the lesion.

No clinical neuronopathic effect has been attributed to Adriamycin in man under current therapeutic regimens. However, neuronal necrosis in peripheral ganglia was seen in patients with leukemia treated with daunorubicin, a closely related drug (12). There is currently an active search for less cardiotoxic analogs of Adriamycin which could be given in higher doses. In view of our finding of more slowly progressing neuronal necrosis in rats given one half the previously determined neurotoxic dose of Adriamycin, one should be alert for similar effects in human subjects who may be given higher doses or more prolonged courses of treatment with Adriamycin or its analogs.

## REFERENCES


1. Blum RH, and Carter SK. Adriamycin. A new anticancer drug with significant clinical activity. Ann Intern Med 80:249–259, 1974.
2. Carter SK. Adriamycin—a review. JNCI 55:1265–1274, 1975.
3. Philips FS, Gilladoga A, Marquardt H, et al. Some observations on the toxicity of Adriamycin (NSC-123127). Cancer Chemother Rep Part 3, vol 6(2):177–181, 1975.
4. Cho ES. Toxic effects of Adriamycin on the ganglia of the peripheral nervous system: a neuropathological study. J Neuropathol Exp Neurol 36:907–915, 1977.
5. Cho ES, Spencer P, Jortner BS, et al. A single intravenous injection of doxorubicin (Adriamycin) induces sensory neuronopathy in rats. Neurotoxicology 1:583–591, 1980.
6. Brown JR. Adriamycin and related anthracycline antibiotics. Prog Med Chem 15:125–164, 1978.
7. Freireich EJ, Gehan EA, Rall DP, et al. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemother Rep 50:219–244, 1966.
8. Jacobs JM, Macfarlane RM, and Cavanagh JB. Vascular leakage in the dorsal root ganglia of the rat, studied with horseradish peroxidase. J Neurol Sci 29:95–107, 1976.
9. Olsson Y. Studies on vascular permeability in peripheral nerves. IV. Distribution of intravenously injected protein tracers in the peripheral nervous system of various species. Acta Neuropathol 17:114–126, 1971.
10. Lee YC, and Byfield JE. Induction of DNA degradation in vivo by Adriamycin. JNCI 57:221–224, 1976.
11. Mahler HR. Nucleic acid metabolism. *In* Basic Neurochemistry, 2nd ed (Siegel GJ, Albers RW, Kazman R, et al, eds). Boston, MA, Little, Brown and Co, 1976.
12. Smith B. Damage to intrinsic cardiac neurones by rubidomycin (daunorubicin). Br Heart J 31:607–609, 1969.

# Skin Rashes Associated With the Administration of the 2-Nitroimidazole, Misonidazole [1]

M. I. Saunders,[2,*] S. Dische,[2] H. D. Kogelnik,[3] Rossall Sealy,[4] and I. Lenox-Smith [5,6]

## SUMMARY

**Misonidazole, a 2-nitroimidazole, is receiving extensive trial as a radiosensitizing drug in the treatment of patients with malignant disease. Drug sensitivity rashes have been observed and the collected cases from three centers are reported. Of a total of 380 patients there were five (1.3%) instances of definite rash and ten (2.6%) instances of observed rash. These figures are compared with the incidence of sensitivity rashes with other drugs in the nitroimidazole group.**

[Cancer Treat Rep 64:263–268, 1980]

The nitroimidazoles became important in medicine > 20 years ago with the introduction of metronidazole (Flagyl), a 5-nitroimidazole which was the first effective trichomonacide in clinical practice. It was soon established as a safe drug with a minimum of side effects. The use of metronidazole has been extended in recent years to the treatment of amebiasis, giardiasis, and infections due to anaerobic bacteria and other organisms. A number of new nitroimidazoles have also been introduced.

In the search for chemical compounds which might restore the radiosensitivity of cells deprived of oxygen, it was discovered that metronidazole was an effective agent (1,2). Improvement in the cure of animal tumors found to contain hypoxic cells was soon demonstrated (3). In the pursuit of more effective radiosensitizing compounds it was found that the Roche experimental drug, misonidazole, was, in laboratory study, more effective than metronidazole (3). The first administration of this drug to man with the object of radiosensitization followed in 1974 (4). Promising observations were made, but a number of toxic effects were observed in subsequent work, including convulsions, peripheral neuropathy, confusional states, nausea, vomiting, and temporary impairment of hearing (5). Despite this, work with this promising radiosensitizing drug has rapidly extended. Presently, many randomized controlled trials are under way and many hundreds of patients have received the drug.

In October 1977, a patient receiving misonidazole manifested a toxic erythema which appeared to be related to the administration of the drug. This was the first observation of skin toxicity due to misonidazole. Subsequently, other cases were observed. This paper records the pooled experience of three centers with regard to skin rashes related to the administration of misonidazole, and the findings recorded by Partington et al (6) in two patients who also showed skin rashes considered to be caused by this drug.

## CASE MATERIAL

The period during which the rashes were observed extends from November 1, 1974, when the first patient received misonidazole at the Mt. Vernon Hospital, to February 28, 1979. During this time, 189 patients at the Mt. Vernon Hospital received between 6 and 48 g of misonidazole in one to 30 doses over 1–48 days. At the Allgemeines Krankenhaus der Stadt Wien, 91 patients received up to 42 g of misonidazole in six to 30 doses over periods of up to 63 days. At the Groote Schuur Hospital and University of Capetown, 100 patients received up to

[1]Received July 25, 1979; revised Nov 30, 1979; accepted Dec 6, 1979.

[2]Marie Curie Research Wing for Oncology, Mt. Vernon Hospital, Northwood, Middlesex, England.

[3]Allgemeines Krankenhaus der Stadt Wien, Universitätsklinik für Strahlentherapie und Strahlenbiologie, Vienna, Austria.

[4]Department of Radiotherapy, Groote Schuur Hospital and University of Capetown, South Africa.

[5]Roche Products Limited, Welwyn Garden City, Hertfordshire, England.

[6]We wish to thank Dr. Dowling Munro (Consultant Dermatologist at St. Bartholomew's Hospital, London, England) for his expert advice, and Audrey Lewis for her preparation of the manuscript.

*Reprint requests to: Dr. M. I. Saunders, Marie Curie Research Wing for Oncology, Regional Radiotherapy Centre, Mt. Vernon Hospital, Northwood, Middlesex HA6 2RN, England.

TABLE 1.—Findings in 12 patients with skin rashes

| Patient No. | Age/ sex | Center | Diagnosis | Date 1st dose of misonidazole given | Total dose of misonidazole before rash appeared (g) | Total No. of doses of misonidazole before rash appeared | Total No. of days over which misonidazole given | Duration of rash | Description of rash |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 44/F | Mt. Vernon Hospital | Carcinoma, cervix | 9/25/77 | 6.5 | 12 | 17 | 7 days | Vesicular rash on palms of hands; a generalized urticarial rash on body and extremities |
| 2 | 65/F | Mt. Vernon Hospital | Carcinoma, bladder | 1/5/78 | 6.5 | 10 | 17 | 10 days | Vesicular rash of hands; macular rash of trunk and extremities |
| 3 | 66/F | Mt. Vernon Hospital | Carcinoma, esophagus | 4/11/78 | 7.0 | 3 | 8 | 13 days | Macular erythematous rash of trunk, with minute vesicles on palms of hands |
| 4 | 47/M | Mt. Vernon Hospital | Recurrent carcinoma, parotid | 7/4/77 | 16.0 | 16 | 22* | 36 days | Macular erythematous rash of trunk and extremities, with vesicular rash on palms of hands and soles of feet |
| 5 | 50/F | Mt. Vernon Hospital | Carcinoma, breast | 11/16/77 | 7.75 | 9 | 13 | 4 days | Macular erythematous rash of trunk and extremities |
| 6 | 62/F | Mt. Vernon Hospital | Carcinoma, breast | 2/16/78 | 8.25 | 11 | 15 | 4 days | Macular erythematous rash of trunk and extremities |
| 7 | 47/F | Mt. Vernon Hospital | Carcinoma, cervix | 5/18/78 | 5.3 | 10 | 15 | 12 days | Macular erythematous rash of trunk and extremities |
| 8 | 53/F | Allgemeines Krankenhaus der Stadt Wien | Recurrent carcinoma, breast; sarcoma, uterus | 4/10/78 | 12.0 | 8 | 9 | About 10 wks | Maculopapular rash of trunk and extremities, not occurring in breast field (5 mev electrons) |
| 9 | 56/M | Groote Schuur Hospital and University of Capetown | Extensive squamous cell carcinoma, mouth | 10/6/77 | 8.0 | 4 | 12 | 2 wks | Deep erythema of trunk and extremities including palms of hands; patchy desquamation of hands |
| 10 | 71/M | Groote Schuur Hospital and University of Capetown | Extensive squamous cell carcinoma, mouth | 8/15/78 | 15.0 | 4 | 10 | Present at 4 days, but fading rapidly | Erythematous macular rash of trunk and extremities, with purpuric element over the shins |
| 11 | 68/F | University of Alberta (Edmonton) | Grade III astrocytoma | 1/16/78 | 12.0 | 6 | 13 | 2 wks | Maculopapular rash of trunk and extremities, with severe pruritis; petechial hemorrhage occurred on the legs when rash was fading |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | 64/M | University of Alberta (Edmonton) | Grade IV astrocytoma | 3/15/78 | 16.0 | 8 | 17 | 7 days | Maculopapular rash of trunk and extremities, with regression of the rash; petechial-like lesions appeared on the legs and abdomen |

*Rash did not develop until 43 days after 1st dose of misonidazole.

TABLE 2.—Results of skin rashes in 12 patients

| Patient No. | Center | Presence of eosinophilia | Further administration of misonidazole | Other drugs | Conclusion |
|---|---|---|---|---|---|
| 1 | Mt. Vernon Hospital | 3200 eosinophils/mm³ on Day 11 | Challenged with misonidazole 4 days after rash disappeared with recurrence of the rash in 8 hrs | None | Toxic erythema due to misonidazole |
| 2 | Mt. Vernon Hospital | 1140 eosinophils/mm³ on Day 6 | No | Oxyprenolol,* thyroxine,* glyceryl trinitrate,* distalgesic* | Erythematous rash due to misonidazole |
| 3 | Mt. Vernon Hospital | Not determined at time of rash | No | Distalgesic | Toxic erythema due to misonidazole |
| 4 | Mt. Vernon Hospital | Not determined at time of rash | Misonidazole was given again 7 mos later without recurrence of rash | None | Toxic vasculitis or erythema multiforme not related to misonidazole |
| 5 | Mt. Vernon Hospital | No eosinophilia | Challenged with misonidazole 19 days later without recurrence of rash; radiotherapy was continued with misonidazole | None | Toxic erythema not due to misonidazole |
| 6 | Mt. Vernon Hospital | Not determined at time of rash | Challenged with misonidazole 5 days later without recurrence of rash; radiotherapy was continued with misonidazole | Librium | Capillaritis probably not related to misonidazole |
| 7 | Mt. Vernon Hospital | 1380 eosinophils/mm³ during time of rash | Radiotherapy was continued with misonidazole with disappearance of rash | Sodium carbenoxolone,* amoxycillin, Valium, distalgesic, disprin and nepenthe, Aludrox, paracetamol, potassium citrate mixture | Toxic erythema not due to misonidazole |
| 8 | Allgemeines Krankenhaus der Stadt Wien | Slight eosinophilia present before misonidazole given; persisted during time of rash | Misonidazole was continued at a daily dose of 1.5 g to a total dose of 18.0 g. Rash progressed with appearance of fever; misonidazole was then stopped | None | Toxic vasculitis due to misonidazole |
| 9 | Groote Schuur Hospital and University of Capetown | Not determined at time of rash | Misonidazole was continued and rash faded; re-administration led to recurrence of rash | None | Erythematous rash due to misonidazole |
| 10 | Groote Schuur Hospital and University of Capetown | No eosinophilia | Misonidazole was discontinued when rash appeared | Theophylline, Panadol, Stelazine, Bactrim | Toxic erythema ? due to misonidazole |
| 11 | University of Alberta (Edmonton) | No eosinophilia | 3 wks later a challenge dose of 500 mg of misonidazole was given with recurrences of rash 10 hrs later | Dexamethasone, diphenylhydantoin, phenobarbital, prochlorperazine | Toxic erythema due to misonidazole |
| 12 | University of Alberta (Edmonton) | No eosinophilia | Challenged with final dose of 2 g of misonidazole with recurrence of rash | Dexamethasone, diphenylhydantoin, prochlorperazine | Toxic erythema due to misonidazole |

*Drugs of long-standing which were continued throughout treatment and unlikely to be significant.

24 g of misonidazole, mainly divided into six doses over 18 days.

The findings in seven patients at the Mt. Vernon Hospital, in one at the Allgemeines Krankenhaus der Stadt Wien, and in two at the Groote Schuur Hospital and University of Capetown are shown in the two summary tables (tables 1 and 2). Two patients from the University of Alberta from the data of Partington et al (6) are also included. Recorded details such as dose and time of administration of misonidazole, appearance and duration of rash, presence of eosinophilia, re-administration of misonidazole, and coincidental administration of other drugs are shown for each patient. With the advice of Dr. Dowling Munro, these data were examined for a relationship of the rash to the administration of the drug. The final conclusions are shown in table 2.

It was concluded that in seven of the cases the rash was certainly related to misonidazole, in one it was possibly related, and in four it was probably not related. Since most of the patients who were given the drug in this early stage of development of this radiosensitizing compound had advanced malignant disease, there were many associated medical problems and a considerable number of other drugs were also given. In these circumstances it is always difficult to be certain of the relationship between the appearance of a skin eruption and the administration of any one drug. In the seven cases where the rash was considered to be definitely related to drug exposure, between 6.5 and 16.0 g was given; the rash appeared between 8 and 17 days after the first dose was given. The rash subsided soon after the drug was discontinued, but was more severe and longer-lasting when the drug was continued.

## DISCUSSION

The original 5-nitroimidazole, metronidazole (Flagyl), has been extensively used worldwide for > 20 years. In a review (7) covering the years 1959–1977, Roe concluded that ". . . skin eruptions have been reported but rarely." This is certainly true, but may indicate under-reporting of an occurrence which is generally not severe or particularly troublesome to the patient. A better idea of the true incidence can be obtained from the collation of results from multi-center trials. Goulton et al (8) found three rashes among 565 patients receiving metronidazole pre- and post-operatively. The first patient had an irritant rash on the face and hands occurring 24 hours after a single 2-g dose; the second patient had a generalized rash appearing on the eighth day of treatment; and the third patient was also receiving cephaloridine so that the origin of the rash is in some doubt.

In another series of studies, Fernex et al (9) reported the results of double-blind trials in 1037 patients receiving either metronidazole or ornidazole, also a 5-nitroimidazole. Skin reactions were seen in three of 469 patients receiving metronidazole and in seven of 568 patients receiving ornidazole. Finally, in a symposium on the treatment of anaerobic infections (10) and in a paper from Greece (11), investigators from five centers reported no skin rashes in 797 patients receiving metronidazole: three of these patients received total doses of 60–100 g. Therefore, the reported incidence of skin rashes with these compounds is between 0 and 1.2%.

The 2-nitroimidazole, benznidazole, has been used in South America for the treatment of Chagas' disease and for leishmaniasis. Barclay et al (12) found skin rashes in 11 of 76 patients receiving benznidazole in increasing doses of 3–7.37 mg/kg/day for 30 days. It is not clear from the paper by Coura et al (13) how many rashes were seen, but side effects were reported in 63% of 309 patients given 5–8 mg/kg/day for 30–60 days. These rashes were dose-related and more common in adults than in children < 15 years of age. Fava et al (14) saw no side effects in six adults given 3–5 mg/kg/day for 45 days. Ferreira (15) observed four rashes in 20 patients given 5–8 mg/kg/day for 62 days (total doses, 21.7–27.8 g). Using a similar dose Levi et al (16) reported five skin rashes in 31 patients. All rashes subsided within three days after the drug was stopped. Thus, in the doses necessary for effective treatment benznidazole is likely to cause skin rashes in 10%–20% of patients, with prolonged treatment being associated with increasing incidence.

Excluding two patients reported by Partington et al (6) where the total treated population is not given, the overall incidence of skin rashes with misonidazole would appear to be five of 380 patients (1.3%), or if all possible rashes relating to these patients are included, ten of 380 patients (2.6%). The potential for misonidazole to cause skin rashes at the doses currently used would seem to be higher with the use of the 5-nitroimidazoles but considerably lower with the use of the other 2-nitroimidazole, benznidazole.

In general, a conservative policy was considered advisable in the use of this new drug, and when associated with a rash it was usually discontinued. If misonidazole is proven beneficial in the treatment of patients with malignant disease, then it might be justified to cautiously continue administration, perhaps under steroid cover. In our patients, treatment of the rash was limited to locally applied steroids and oral administration of antihistamines. The rash soon subsided without complication in all but one of these patients (No. 8).

It is important that those investigators now using misonidazole recognize that skin rashes may occur in their patients. This complication has not proved to be a serious one, but patients presenting with such rashes must usually discontinue taking the drug. The frequency of its occurrence is small, and the skin eruptions are not likely to significantly affect the results of the controlled trials currently under way or the future use of the drug if it is found to be a valuable radiosensitizer for clinical use.

## REFERENCES


1. Adams GE, and Dewey DL. Hydrated electrons and radiobiological sensitization. Biochem Biophys Res Commun 12:473–477, 1963.
2. Asquith JC, Foster JL, Willson RL, et al. Metronidazole (Flagyl). A radiosensitizer of hypoxic cells. Br J Radiol 47:474–481, 1974.
3. Fowler JF, Adams GE, and Denekamp J. Radiosensitizers of hypoxic cells in solid tumours. Cancer Treat Rev 3:227–256, 1976.
4. Gray AJ, Dische S, Adams GE, et al. Clinical testing of the radiosensitizer Ro 07-0582. I. Dose tolerance, serum and tumour concentration. Clin Radiol 27:151–157, 1976.
5. Dische S, Saunders MI, Flockhart IR, et al. Misonidazole—a drug for trial in radiotherapy and oncology. Int J Radiat Oncol Biol Phys 5:851–860, 1979.
6. Partington J, Koziol D, Chapman D, et al. A new side effect of the hypoxic cell sensitizer, misonidazole. Cancer Treat Rep 63:123–125, 1979.
7. Roe JFC. Metronidazole: review of uses and toxicity. J Antimicrob Chemother 3:205–212, 1977.
8. Goulton J, Wilkinson MA, and Baker PG. Multicentre study of Flagyl in the prevention of post-operative anaerobic infection. J Int Med Res 7:100–105, 1979.
9. Fernex M, Jeunet F, and Richle R. Development of a nitroimidazole derivative (Ro 07-0207) for the treatment of amoebiasis, lambliasis and trichomoniasis. *In* Progress in Chemotherapy (Daikos GK, ed). Athens, Greece, Hellenic Society of Chemotherapy, 1974, vol 1, pp 978–982.
10. Willis AT, Leigh DA, and Hamilton-Miller J, eds. Management of non-clostridial anaerobic infections. J Antimicrob Chemother (Suppl C)4:25–81, 1978.
11. Giamarellou H, Kanellakopoulou K, Pragastis D, et al. Treatment with metronidazole of 48 patients with serious anaerobic infections. J Antimicrob Chemother 3:347–353, 1977.
12. Barclay CA, Cerisola JA, Lugones H, et al. Status of the clinical and seroparasitological evaluation of benznidazole in the treatment of acute Chagas' disease. *In* Current Chemotherapy (Siegenthaler W, and Luthy R, eds). Washington, DC, American Society of Microbiology, 1978, vol 1, pp 158–159.
13. Coura JR, Brindeiro PJ, and Ferreira I. Benznidazole in the treatment of Chagas' disease. Washington, DC, American Society of Microbiology, 1978, vol 1, pp 161–162.
14. Fava S di C, Zamith VA, Cuce LC, et al. Treatment of American mucocutaneous leishmaniasis with benznidazole. Washington, DC, American Society of Microbiology, 1978, vol 1, pp 163–164.
15. Ferreira HD. Treatment of Chagas' disease with Ro 07-1051. Rev Inst Med Trop Sao Paulo 18:357–364, 1976.
16. Levi GC, Amato Neto V, and de Araújo Bezerra Sant'anna IF. Side effects of Ro 07-1051, a nitroimidazole used in the experimental treatment of Chagas' disease. Rev Inst Med Trop Sao Paulo 17:49–52, 1975.

# Salvage Treatment of Patients Relapsing After Breast Cancer Adjuvant Chemotherapy [1,2]

**Albert G. Wendt, Stephen E. Jones,* and Sydney E. Salmon [3,4]**

## SUMMARY

**We analyzed the efficacy of salvage therapy with systemic agents in 33 women who initially received adjuvant chemotherapy with Adriamycin and cyclophosphamide (AC) for early breast cancer. Relapses occurred at a median of 9 months after completion of adjuvant therapy, and six patients relapsed while receiving adjuvant treatment. Nonetheless, salvage treatment produced objective responses in three of ten patients (30%) receiving hormonal therapy alone, in five of 11 (45%) receiving AC plus hormonal therapy, and in none of 12 receiving cyclophosphamide, methotrexate, and 5-fluorouracil ($P < 0.05$). Survival from the time of relapse from adjuvant chemotherapy was superior for patients receiving hormonal therapy with or without AC chemotherapy compared to that of patients receiving cyclophosphamide, methotrexate, and 5-fluorouracil ($P < 0.05$), and the median survival time of responders is > 12 months. On the basis of these findings, we believe that patients relapsing after adjuvant chemotherapy should receive aggressive treatment employing Adriamycin combination chemotherapy along with hormonal therapy.**

[Cancer Treat Rep 64:269–273, 1980]

Adjuvant chemotherapy after primary surgical treatment of breast cancer has been clearly shown to be of benefit to premenopausal patients (1,2). Recently, there has also been evidence that full doses of cyclophosphamide, methotrexate, and 5-fluorouracil (5-FU) (CMF) or Adriamycin-containing regimens produce a similar benefit in disease-free survival in postmenopausal patients (3–6). The concept of initiating early treatment of micrometastases (when the tumor burden is low) rather than waiting for the development of overt metastases is rapidly becoming the standard of care in breast cancer. This approach is reasonable in view of the positive results in early breast cancer and the current inability to eradicate overt metastatic breast cancer with currently available anticancer drugs or hormones. Despite the efficacy of adjuvant therapy for many patients, a proportion of patients receiving adjuvant chemotherapy still recur with currently available treatment, and the oncologist is increasingly faced with the problem of planning therapy for adjuvant treatment failures. Little information is currently available on this topic. Accordingly, we have reviewed our experience with the management of 33 patients who recurred after receiving initial adjuvant chemotherapy with Adriamycin and cyclophosphamide (AC).

## METHODS AND PATIENTS

Since July 1974, > 250 patients with primary breast cancer have been treated with AC in an adjuvant chemotherapy program at the University of Arizona. The details and results of the study have been published elsewhere (4,7,8). Adjuvant chemotherapy consists of Adriamycin given at a dose of 30 mg/m$^2$ iv on Day 1 followed by cyclophosphamide given at a dose of 150 mg/

[1]Received July 31, 1979; revised Nov 28, 1979; accepted Dec 10, 1979.

[2]Supported in part by Public Health Service grants CA-17094 and CA-23074 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Section of Hematology and Oncology, University of Arizona College of Medicine, Tucson.

[4]We thank Thomas E. Moon, PhD, for statistical assistance and Toni Meinke for secretarial assistance.

**Reprint requests to:* Stephen E. Jones, MD, Section of Hematology and Oncology, The University of Arizona Health Sciences Center, 1501 N Campbell Ave, Tucson, AZ 85724.

TABLE 1.—Comparison of clinical features and sites of involvement at relapse after adjuvant treatment

| Clinical feature | Therapy | | | Significance* |
|---|---|---|---|---|
| | CMF (12 patients) | Hormonal (10 patients) | AC + hormonal (11 patients) | |
| Median time to relapse | 5 mos | 18 mos | 12 mos | NS |
| No. of patients with— | | | | |
| Menopausal status | | | | |
| Premenopausal | 5 | 4 | 4 | NS |
| Postmenopausal | 7 | 6 | 7 | |
| Estrogen-receptor status | | | | |
| Positive | 1 | 1 | — | NS |
| Negative | 2 | — | — | |
| Sites of involvement | | | | NS |
| Soft tissue | 2 | 1 | 3 | |
| Lung | 3 | 1 | 3 | |
| Nodes | 4 | 2 | 0 | |
| Bone | 6 | 6 | 5 | |
| Liver | 2 | 1 | 1 | |
| Pleura | 0 | 2 | 2 | |
| Bone marrow | 1 | 1 | 0 | |
| Brain | 0 | 1 | 0 | |
| Ovary | 0 | 1 | 0 | |
| Stage IV NED† | 0 | 1 | 1 | |
| Radiotherapy for painful bony metastases | 2 | 5 | 2 | |

*NS = not significant.
†Stage IV NED = patients with no evidence of disease whose local recurrence was removed with biopsy.

$m^2$ orally on Days 3–6. Treatment is repeated every 3 weeks for a total of eight cycles for patients with positive axillary lymph nodes (stages II and III). Patients with negative nodes (stage I) receive only three courses of AC. The usual total dose of Adriamycin administered, therefore, is 240 $mg/m^2$, thus leaving considerable leeway for future Adriamycin administration, if needed.

Thirty-three patients have developed metastatic breast cancer after receiving adjuvant AC chemotherapy. The distribution of these patients by initial extent of disease was: *stage I*—negative nodes, two patients; *stage II*—one to three positive nodes, six patients and four or more positive nodes, 11 patients; and *stage III*—14 patients. The type of initial salvage treatment, response to treatment, and survival were reviewed. During this time, there was no standard approach to management and treatment was individualized. For purposes of analysis, the type of therapy could be divided into three groups: combination chemotherapy with CMF (9), hormonal treatment alone, and combination chemotherapy with AC plus hormonal therapy. The features of the 33 patients are shown in table 1. Nine patients with painful bony lesions at the time of relapse received palliative local radiotherapy. All patients except two had measurable (nonirradiated) tumor by which response could be judged. All patients were carefully assessed for metastatic sites at the time of relapse (eg, chest radiographs, bone and liver scans, etc), and the patient characteristics in the three treatment groups were similar (table 1). Response to systemic salvage therapy was judged as complete (disappearance of all known disease), partial (reduction by one half of all measurable disease), symptomatic improvement (without objective response), stable disease, or progressive disease. Two patients were considered to be stable, without evidence of disease if all known disease was excised at the time of relapse in confirming the diagnosis, eg, a skin or regional lymph node metastasis.

Of those patients treated with standard-dose CMF (9), seven received CMF alone and five others received CMF along with additional treatment: Halotestin (one), vincristine (one), hexamethylmelamine, prednisone, and megestrol (two), and X-ray ablation of the ovaries plus megestrol (one).

In the hormonal therapy-alone group, treatment consisted of tamoxifen (six patients), tamoxifen and oophorectomy (one), diethylstilbestrol (one), Halotestin (one), and oophorectomy plus adrenalectomy (one).

In those patients to whom AC was given, concurrent hormonal therapy was offered in all but one case; it consisted of oophorectomy (two patients), oophorectomy and

TABLE 2.—Results of salvage treatment in 33 women with recurrent breast cancer after relapse from adjuvant AC

| Treatment group | No. of patients | | | | | |
|---|---|---|---|---|---|---|
| | Response | | | | | |
| | Total | Complete* | Partial* | Symptomatic | Stable | Progression |
| CMF | 12 | 0 | 0 | 2 | 1 | 9 |
| Hormonal therapy | 10 | 1 | 2 | 0 | 3 | 4 |
| AC + hormonal therapy | 11 | 2 | 3 | 1 | 3 | 2 |
| Totals | 33 | 3 (9%) | 5 (15%) | 3 | 7 | 15 |

*Objective response rate of 38% for the 21 patients receiving hormonal therapy and AC + hormonal therapy was significantly greater than that observed for CMF ($P = < 0.05$) since there were no responders in the latter group.

TABLE 3.—Characteristics of patients responsive to salvage treatment

| Patient No. | Age (yrs) | Menopausal status | Interval to relapse (mos)* | Sites of relapse | Relapse treatment | Response | Response duration (mos) |
|---|---|---|---|---|---|---|---|
| | | | | *Hormonal therapy* | | | |
| 1 | 51 | Postmenopausal | 20 | Bone marrow, ovary, liver | Oophorectomy, adrenalectomy | Partial | 10 |
| 2 | 60 | Postmenopausal | 6 | Bone, pleura | Diethylstilbestrol | Complete | 24+ |
| 3 | 56 | Postmenopausal | 16 | Skin, pleura | Tamoxifen | Partial | 11 |
| | | | | *AC + hormonal therapy* | | | |
| 4 | 56 | Postmenopausal | Receiving AC | Pleura | Tamoxifen, AC + vincristine | Complete | 6+ |
| 5 | 55 | Postmenopausal | 28 | Lung, bone | Tamoxifen, AC, prednisone, megestrol | Partial | 14 |
| 6 | 69 | Postmenopausal | Receiving AC | Bone | Tamoxifen + AC + vincristine | Partial | 6+ |
| 7 | 39 | Premenopausal | 11 | Lung, skin | Tamoxifen, AC, prednisone, megestrol, oophorectomy | Partial | 4 |
| 8 | 72 | Postmenopausal | 18 | Liver | Tamoxifen + AC | Complete | 12+ |

*Interval from the completion of adjuvant treatment to relapse.

megestrol (one), oophorectomy, megestrol, and prednisone (one), tamoxifen (four), megestrol and prednisone (one), and tamoxifen and prednisone (one). AC was given with vincristine to six patients and with 5-FU to two patients. Doses of AC were as previously described (Adriamycin at a dose of 40 mg/m$^2$ on Day 1 and cyclophosphamide at a dose of 200 mg/m$^2$ daily orally for 4 days [Days 3–6]) (10).

## RESULTS

The results of initial relapse treatment are summarized in table 2. In these 33 patients, there were three complete responses (9%) and five partial responses (15%) to salvage treatment, for an overall objective response rate of 24%. There were, however, no objective responses to treatment with CMF combinations and CMF was significantly less effective than hormonal therapy (30% response rate) or AC plus hormonal therapy (45% response rate) ($P < 0.05$). Treatment, the type and duration of response, and patient characteristics of the responding patients are shown in table 3.

Three additional patients who were given hormonal treatment were stable without evidence of disease for periods of 6–8 months. One patient receiving CMF was stable without evidence of disease for 3 months, while two other similarly treated patients had symptomatic but not objective improvement. In those patients receiving AC plus hormonal therapy, there was one who experienced symptomatic improvement and three who were stable for periods of 4–12 months.

Figure 1 is an actuarial plot of the survival from time of adjuvant relapse of the 33 patients according to initial salvage treatment. There was no statistically significant difference in survival time from relapse between hormonal therapy and hormonal therapy plus AC combinations. Both treatment groups, however, were superior to the CMF group ($P = 0.004$, $P = 0.01$). The median survival time after relapse was only 6 months in the CMF group and has not yet been reached in the AC plus hormonal therapy group.

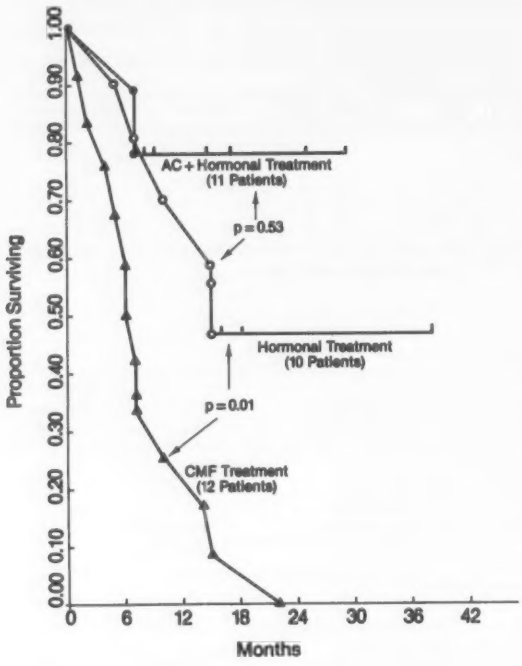



FIGURE 1.—Survival of 33 patients from time of relapse after adjuvant chemotherapy according to type of initial relapse therapy.

For the 33 patients, the median time to relapse from initial adjuvant therapy was 9 months after completion of treatment, with the last relapse occurring 33 months after therapy. Six patients relapsed while still receiving adjuvant chemotherapy. There was no significant difference between times to relapse for the three relapse treatment groups, although the median time was somewhat longer in the hormonal therapy-alone group (18 months) compared to the AC plus hormonal therapy (12 months) or CMF (5 months) group.

For all patients, those with objective response (partial or complete) to salvage treatment survived significantly longer (median, > 12 months) than nonresponders (median, 6 months) ($P = 0.01$). The median survival from time of relapse was 6 months for patients receiving CMF salvage treatment (none of whom achieved an objective response) (fig 1).

## DISCUSSION

The purpose of this analysis was to obtain some assessment of the relative effectiveness of salvage treatment for breast cancer recurrent after adjuvant treatment with AC. Estrogen-receptor information was not available in the majority of our patients. Nonetheless, both hormonal therapy alone or re-induction chemotherapy with AC plus hormonal treatment produced worthwhile remissions. Hormonal therapy alone produced objective responses (including one complete) in 30% of patients. AC, when used in combination with hormonal therapy, produced a 45% objective response rate despite previous adjuvant use of these two drugs. However, because the number of patients is small, this is not significantly different than hormonal therapy alone. More importantly, the use of CMF appeared to have little efficacy in this setting with not a single objective response observed among 12 patients. The patients receiving CMF salvage treatment had a median survival time after adjuvant relapse of only 6 months (fig 1). It is possible that the use of tamoxifen or other hormonal treatment with CMF might improve the results achievable with this combination.

One difference between treatment groups in our series was the interval to relapse after completion of adjuvant treatment. This was shorter in the CMF group (although not significantly so) and might suggest that this group consisted of patients with more aggressive disease, less likely to respond to chemotherapy, who failed early after adjuvant treatment. There were, however, three patients who relapsed ≥ 1 year after completion of adjuvant therapy who failed to respond to salvage treatment with CMF combinations. Of interest, objective responses occurred in two of six patients who had relapsed while still receiving their initial adjuvant treatment. Both of these patients received intensified AC chemotherapy with hormonal therapy, demonstrating that early relapse, even during therapy, does not preclude response to more intensive salvage treatment.

Analysis of the survival curves after relapse from initial adjuvant therapy demonstrates that patients treated with hormonal therapy, with or without AC combination chemotherapy, and patients with objective responses survive significantly longer than those treated with CMF or those without objective response. Thus, in addition to producing objective response, effective salvage treatment translates into direct survival benefit.

There is little published information on the management of adjuvant failures. In the Milan series where patients received 12 months of adjuvant chemotherapy with CMF, Bonadonna et al (1) reported the initial use of endocrine treatment in 21 patients and combination chemotherapy in 33 patients (primarily Adriamycin and vincristine) (1). The projected median duration of survival was about 2 years for both of these groups of patients. No data were provided on the frequency of objective responses to initial relapse therapy (1).

Although the numbers of patients in this analysis are small, we have now adopted the following approach to patients who relapse after adjuvant chemotherapy: (*a*)

we do not use CMF and (*b*) we recommend aggressive treatment of first recurrence with hormonal measures (additive or ablative) along with full-dose Adriamycin combination chemotherapy. In a parallel series of studies for patients with advanced breast cancer (but no prior cytotoxic chemotherapy), we found that the addition of either vincristine or the androgen, calusterone, to AC produced prolongation of relapse-free survival and overall survival compared to the use of AC alone (11,12). Thus, our present program for adjuvant failures includes the use of tamoxifen (since calusterone is not presently available and tamoxifen is probably a more active hormonal agent than estrogens or androgens [13]) along with AC and vincristine (TACO).

## REFERENCES


1. Bonadonna G, Valagussa P, Rossi A, et al. Are surgical adjuvant trials altering the course of breast cancer? Semin Oncol 5:450–464, 1978.
2. Fisher B, and Redmond C. Studies of the National Surgical Adjuvant Breast Project (NSABP). *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 67–81.
3. Bonadonna G, Valagussa P, Rossi A, et al. CMF adjuvant chemotherapy in operable breast cancer. *In* Adjuvant Therapy of Cancer II (Jones SE, and Salmon SE, eds). New York, NY, Grune and Stratton, 1979, pp 227–236.
4. Wendt AG, Jones SE, Salmon SE, et al. Adjuvant treatment of breast cancer with Adriamycin/cyclophosphamide with or without radiation therapy. *In* Adjuvant Therapy of Cancer II (Jones SE, and Salmon SE, eds). New York, NY, Grune and Stratton, 1979, pp 285–293.
5. Williams SD, and Einhorn LH. Adriamycin chemoprophylaxis of high risk breast cancer. *In* Adjuvant Therapy of Cancer II (Jones SE, and Salmon SE, eds). New York, NY, Grune and Stratton, 1979, pp 295–302.
6. Buzdar A, Blumenschein G, Gutterman J, et al. Adjuvant therapy with 5-fluorouracil, Adriamycin, cyclophosphamide and BCG (FAC-BCG) for stage II or III breast cancer. *In* Adjuvant Therapy of Cancer II (Jones SE, and Salmon SE, eds). New York, NY, Grune and Stratton, 1979, pp 277–284.
7. Hammond N, Jones SE, Salmon SE, et al. Adjuvant treatment of breast cancer with Adriamycin-cyclophosphamide with or without radiation therapy. *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 153–160.
8. Salmon SE, Wendt AG, Jones SE, et al. Treatment of early breast cancer with Adriamycin-cyclophosphamide with or without radiation therapy: initial results of a brief and effective adjuvant program. Recent Results Cancer Res 68:98–104, 1979.
9. Canellos GP, DeVita VT, Gold LG, et al. Cyclical combination chemotherapy for advanced breast carcinoma. Br Med J 1:218–220, 1974.
10. Jones SE, Durie BGM, and Salmon SE. Combination chemotherapy with Adriamycin and cyclophosphamide for advanced breast cancer. Cancer 36:90–97, 1975.
11. Rainey JM, Jones SE, and Salmon SE. Combination chemotherapy for advanced breast cancer utilizing vincristine, Adriamycin, and cyclophosphamide (VAC). Cancer 43:66–71, 1979.
12. Lloyd RE, Jones SE, and Salmon SE. Comparative trial of low-dose Adriamycin plus cyclophosphamide with or without additive hormonal therapy in advanced breast cancer. Cancer 43:60–65, 1979.
13. Legha SS, Buzdar AU, Hortobagyi GN, et al. Tamoxifen. JAMA 242:49–52, 1979.

# Effect of Hepatic Microsomal Enzyme Inducers on the Metabolism of Misonidazole in Rats [1]

Dale Shoemaker,* Debbie Upton, and John Strong [2,3]

## SUMMARY

**Pretreatment of rats with the hepatic microsomal enzyme inducers, phenobarbital or phenytoin, increased the metabolism of the hypoxic cell radiosensitizer, misonidazole. The increase was reflected in the ratio of *O*-demethylated metabolite to unchanged drug in the 24-hour urine collection and was more extensive with phenobarbital. No inducement was observed after pretreatment with dexamethasone or misonidazole. Such pretreatment with inducers prior to misonidazole therapy would be clinically important since the incidence of neurotoxicity may be reduced without impairment of the radiosensitizing effect.**

[Cancer Treat Rep 64:275–277, 1980]

Misonidazole, a hypoxic cell radiosensitizer, is presently undergoing clinical trials. Two recent papers (1)[4] have reported that the incidence of misonidazole-associated neurotoxicity observed during these trials is less in patients receiving continuous therapy with drugs such as phenytoin which induce hepatic microsomal enzymes. After the iv administration of misonidazole, patients treated continuously with phenytoin exhibited a much shorter plasma half-life, 50% of that of patients receiving a similar dose schedule with no additional drugs. The peak plasma levels, however, were similar in both groups. The ratio of metabolite to parent drug excreted in the urine over a 24-hour period was approximately sixfold greater in a patient receiving phenytoin as compared to the control group, suggesting an increased metabolism of misonidazole. Since the time of exposure and the dose of misonidazole are important factors in the incidence of neurotoxicity (2), pretreatment with inducers of hepatic enzymes which markedly shorten the plasma half-life would be clinically significant. This is especially important when coupled with the iv route of administration to reduce gastrointestinal disturbances and the proper timing of radiation therapy at the maximum tissue concentration. Therefore, the effect of hepatic enzyme induction on misonidazole metabolism, as reflected in the ratio of metabolite to parent in the urine, was determined by the pretreatment of rats with one of two known inducers, phenobarbital or phenytoin. Animals were also pretreated with misonidazole and dexamethasone to determine if induction was possible from prolonged treatment with the radiosensitizer or from continuous treatment with a drug frequently given to patients receiving misonidazole respectively.

## MATERIALS AND METHODS

Male Sprague-Dawley rats weighing 175–200 g were pretreated with 0.9% saline (0.25 ml ip for 7, 12, or 18 days), misonidazole (100 mg/kg ip for 7 days), phenobarbital sodium (75 mg/kg ip for 7 days), dexamethasone (0.23 mg/kg ip for 7 days), or phenytoin sodium (~ 50 mg/kg ip for 18 days; 100 mg/kg ip for 12 days; and 200 mg/kg orally for 12 days). Twenty-four hours after the last pretreatment, $^{14}$C-misonidazole (0.5 ml/rat of a 24-mg/ml dose in 0.9% saline solutions, 3 microcuries/rat) was administered via the femoral vein. The animals were placed in metabolic cages and the urine and feces were collected

[1]Received July 31, 1979; revised Dec 13, 1979; accepted Dec 19, 1979.

[2]Laboratory of Chemical Pharmacology, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD.

[3]We thank Richard Whyte for his technical assistance.

[4]Wasserman TH, Phillips TL, Van Raalte G, et al. The neurotoxicity of misonidazole: potential modifying role of phenytoin sodium and dexamethasone. Manuscript submitted to *British Journal of Radiology*.

**Reprint requests to:* Dale Shoemaker, PhD, Laboratory of Chemical Pharmacology, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bldg 37, Rm 5A19, National Institutes of Health, Bethesda, MD 20205.

TABLE 1.—Parameters of misonidazole metabolism in normal and pretreated rats

| Treatment group | No. of rats treated | Dose and route | 24-hr urine concentration (mg) of— | | Ratio of metabolite to misonidazole | % dose in urine | % dose in feces | Ratio of feces to urine |
|---|---|---|---|---|---|---|---|---|
| | | | Misonidazole | Metabolite | | | | |
| Saline | 6 | 0.25 ml ip | 1.4* | 1.0* | 0.8 | 39.4* | 27.8* | 0.71 |
| Misonidazole | 3 | 100 mg/kg ip | 1.6 | 1.1 | 0.7 | 47.3 | 36.7 | 0.78 |
| Phenobarbital | 4 | 75 mg/kg ip | 0.6 | 1.5 | 2.5 | 42.9 | 30.6 | 0.71 |
| Dexamethasone | 4 | 0.23 mg/kg ip | 1.8 | 1.1 | 0.6 | 40.3 | 38.6 | 0.96 |
| Phenytoin | 4 | ~ 50 mg/kg ip | 1.3 | 1.0 | 0.8 | 40.0 | 29.7 | 0.74 |
| Phenytoin | 3 | 100 mg/kg ip | 1.3 | 1.4 | 1.1 | 44.7 | 21.4 | 0.48 |
| Phenytoin | 4 | 200 mg/kg orally | 1.2 | 1.4 | 1.2 | 48.4 | 20.4 | 0.42 |

*Each value = mean value for No. of rats treated.

for 24 hours. Concentrations of misonidazole and the *O*-demethylated metabolite, desmethylmisonidazole, in the excreta were determined by a modification of the high-performance liquid chromatography (HPLC) method of Workman et al (3). The percent dose of $^{14}$C-misonidazole in the urine and feces was found by liquid scintillation counting in a Searle Mark III counter. Aliquots of urine samples were added directly to scintillation fluid and counted. The feces were homogenized in 4 vol of distilled water and aliquots were digested with NCS solubilizer prior to the addition of scintillation fluid and counting. Samples were corrected for quench using an automatic external standard.

To examine for the possible conjugation of misonidazole and the *O*-demethylated metabolite, 0.5-ml aliquots of fecal homogenates were extracted twice with 2 ml of methanol, dried under vacuum, and hydrolyzed with either β-glucuronidase (type VIII) for 24 hours at 37°C (4000 units in 0.1 M acetate, pH 5.0) or 6 N HCl for 2 hours at 95°C–100°C. The reaction mixtures were analyzed for parent and metabolite by the HPLC method as mentioned above. Reference misonidazole and *O*-demethylated metabolite samples were subjected to hydrolysis procedures to establish the extent of possible degradation of the compounds under these conditions.

## RESULTS

The concentrations of misonidazole and the *O*-demethylated metabolite in 24-hour urine samples of rats pretreated with known hepatic microsomal enzyme inducers indicated that an extensive increase in metabolism had occurred (table 1). In rats pretreated with phenobarbital, the amount of misonidazole decreased by > 50%, the amount of metabolite increased by approximately 40%, and the metabolite to parent drug ratio increased by 250% as compared to controls. The values for the saline-pretreated control group closely agreed to those of normal rats reported by Flockhart et al (4). The percentages of the dose recovered as misonidazole and the metabolite in the urine were 10.0% and 9.0% respectively versus 11.6% and 9.1% in the Flockhart study. The ratio of metabolite to parent drug recovered in the urine was identical, 0.9, while the percent recovery of total $^{14}$C was 67% in this study versus 87% in the Flockhart study. The latter difference was possibly due to a longer fecal collection period. When rats were pretreated with phenytoin, a less potent inducer than phenobarbital (5), a similar pattern of parent drug and metabolite concentrations and metabolite to parent drug ratio as that noted above was observed, although the percentages of increase and decrease were smaller. The induction with phenytoin appeared to be dependent on the dose since no increase in metabolism occurred at an ip dose of approximately 50 mg/kg, but was seen at an ip dose of 100 mg/kg. A corresponding increase in induction to the latter also resulted from oral pretreatment at 200 mg/kg. It should be noted also, that the latter pretreatment doses of 100 mg/kg ip and 200 mg/kg orally were made for 12 days prior to $^{14}$C-misonidazole versus 18 days for the dose of 50 mg/kg ip, pointing out again that the dependence of induction appears to be due to dose rather than to the number of days of pretreatment. Based on a previous report suggesting self-induction in nonhuman primates (6), one group of rats was pretreated with misonidazole. The induction, as reflected by metabolism, was observed to be essentially the same as that of the saline controls. Dexamethasone was used in the pretreatment since patients receiving continuous dexamethasone and/or phenytoin therapy have shown a decrease in the plasma half-life of misonidazole. There appeared to be no induction and animals in this pretreatment group had a significant loss of body weight.

The ratio of percent dose as $^{14}$C in the feces to percent dose in the urine was constant in the saline, misonidazole, low-dose ip phenytoin, and phenobarbital treatment groups (table 1). However, this ratio was increased in the dexamethasone group while it was decreased in the higher-dose phenytoin groups. After pretreatment with inducers, very minimal amounts of unchanged drug or metabolite were found to be present in the feces. Analysis by HPLC of fecal homogenates subjected to β-glucuron-

idase incubation or HCl hydrolysis indicated that in all cases the concentration of the parent or metabolite moieties did not increase and was < 0.5% of the administered dose. Misonidazole and metabolite standards were used as controls in the HCl hydrolysis. The treatment-associated decrease in their concentration was negligible.

## DISCUSSION

After pretreatment with phenobarbital and phenytoin, the ratio of metabolite to unchanged drug in the urine was increased as compared to controls and appeared to result from an induction of the hepatic microsomal enzymes. In phenobarbital-pretreated rats the decrease in parent drug concentration was not totally accounted for by the increase in metabolite, indicating that other routes of metabolism were possibly induced. However, if other paths were induced, they had no effect on the ratio of fecal to urine excretion. In the higher-dose phenytoin-pretreated rats the decrease of unchanged drug in the urine was essentially balanced by the increase in concentration of the metabolite.

The enhanced metabolism seen after pretreatment with inducers was reflected in this study by the increased urine metabolite to unchanged drug ratios and in other studies (1)[5] by the decreased plasma half-life and shorter times of tissue exposure. These effects would not be expected to impair the sensitizing action of misonidazole if the irradiation is properly timed to peak plasma levels since phenytoin pretreatment does not significantly alter peak levels when compared to controls (1). Therefore, the incidence of neurotoxicity which is related to the time of exposure and the dose of misonidazole (2) may be lessened by treatment with inducers prior to misonidazole administration.

[5]*See* footnote 4.

## REFERENCES


1. Gangji D, Schwade JG, and Strong JM. Phenytoin-misonidazole: possible metabolic interaction. Cancer Treat Rep 64:155–156, 1980.
2. Dische S, Saunders MI, Lee ME, et al. Clinical testing of the radiosensitizer Ro 07-0582: experience with multiple doses. Br J Cancer 35:567–573, 1977.
3. Workman P, Little CJ, Marten TR, et al. Estimation of the hypoxic cell-sensitiser misonidazole and its *O*-demethylated metabolite in biological materials by reversed-phase high performance liquid chromatography. J Chromatogr 145:507–512, 1978.
4. Flockhart IR, Large P, Troup D, et al. Pharmacokinetic and metabolite studies of the hypoxic cell radiosensitizer misonidazole. Xenobiotica 8:97–105, 1978.
5. Kato R, and Vassanelli P. Induction of increased meprobamate metabolism in rats pretreated with some neurotropic drugs. Biochem Pharmacol 11:779–794, 1962.
6. Johnson R, Gomer C, Ambrus J, et al. An investigation of the pharmacological and radiosensitizing effects of the 2-nitroimidazole Ro-07-0582 in primates. Br J Radiol 49:294, 1976.

# Vinblastine Given as a Continuous 5-Day Infusion in the Treatment of Refractory Advanced Breast Cancer[1,2]

H.-Y. Yap,* G. R. Blumenschein, M. J. Keating, G. N. Hortobagyi, C. K. Tashima, and T. L. Loo[3]

## SUMMARY

**In an attempt to improve the therapeutic index of vinblastine by maintaining a prolonged, constant plasma level, it was given as a continuous 5-day infusion at 1.4–2.0 mg/m²/day. Among the 30 evaluable patients with advanced refractory breast cancer who had been extensively treated previously, there were one complete and 11 partial responders, with an overall response rate of 40%. Six of the 12 patients who had previously received a conventional dose schedule of vinca alkaloids responded, including two who had progressed with intermittent bolus iv injections of vinblastine. Responses were seen in all organ sites of involvement, with durations of 2–8+ months (median, 6). Myelosuppression was mild to moderate at doses of between 1.4 and 1.8 mg/m²/day but became severe at 2 mg/m², with leukopenia (wbcs, < 2000/mm³) and granulocytopenia (absolute granulocytes, < 1000/mm³) occurring in 75% of administered courses. Other toxic effects were infrequent and acceptable. Our results suggest that vinblastine given as a continuous 5-day infusion is highly effective in the treatment of patients with advanced metastatic breast cancer, including a few whose disease was felt to be resistant to a conventional dose schedule of vinca alkaloids.**

[Cancer Treat Rep 64:279–283, 1980]

The introduction of combination chemotherapy has led to significant improvement in the response and survival rate of patients with metastatic breast cancer (1–4). We previously reported that a combination of 5-fluorouracil (5-FU), Adriamycin, and cyclophosphamide, with or without immunotherapy, has resulted in an overall response rate of 70%, with a median duration of 14–16 months (5–8). Nonetheless, a small proportion of patients have failed to respond to 5-FU, Adriamycin, and cyclophosphamide combination chemoimmunotherapy, and almost all patients, regardless of their initial response, will ultimately relapse and die from their disease. There is, therefore, a need to explore not only new chemotherapeutic agents, but also dose schedules with drugs of established antitumor activity.

Vinblastine, a mitotic inhibitor, has been reported to have moderate activity in patients with metastatic breast cancer. A review of the literature (9–13) revealed a response rate of about 20% when vinblastine was given as an intermittent iv bolus injection. Pharmacokinetic studies have demonstrated that, following single iv administration, the drug is rapidly cleared from the blood in a triphasic manner, with an initial half-life of 0.07 hour and a terminal half-life of about 24 hours (14,15). It was possible that a continuous infusion of vinblastine might improve its therapeutic index by maintaining a prolonged, constant plasma level. Furthermore, preliminary clinical studies have suggested that vindesine (with similar pharmacokinetic parameters to vinblastine) is more effective when given as a continuous iv infusion (16). We therefore investigated a continuous 5-day infusion schedule of vinblastine in patients with advanced refractory metastatic breast cancer.

[1]Received Aug 24, 1979; revised Nov 26, 1979; accepted Dec 5, 1979.

[2]Presented in part at the American Society of Clinical Oncology Meeting, New Orleans, LA, May 1979.

[3]Medical Breast Service, Department of Medicine (*H.-Y. Yap, G. R. Blumenschein, G. N. Hortobagyi,* and *C. K. Tashima*), and Department of Developmental Therapeutics (*M. J. Keating* and *T. L. Loo*), The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston.

**Reprint requests to:* H.-Y. Yap, MD, Department of Medicine, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, 6723 Bertner Ave, Houston, TX 77030.

## MATERIALS AND METHODS

Thirty-four patients with histologically proven breast cancer refractory to more conventional chemotherapeutic agents, including Adriamycin, 5-FU, cyclophosphamide, and methotrexate, were entered in this study. The starting dose for the first 15 patients was 2.0 mg/m²/day given as a continuous steady-rate infusion for 5 days and repeated every 3 weeks. Because of the excessive myelotoxicity observed at the 2-mg/m² dose level, we reduced the dose to 1.6 mg/m²/day for 5 days every 3 weeks. To avoid the likelihood of severe tissue necrosis, central venous delivery was utilized. Indwelling silicone elastomer venous catheters were inserted peripherally via an arm vein in six patients and via infraclavicular sc catheterization of the subclavian vein in the remaining 28 patients. To ensure a constant infusion rate, the total daily dose of vinblastine was dissolved in 1000 ml of 5% dextrose in water given as a continuous 24-hour infusion via the McGraw volumetric infusion (inpatient). Towards the latter part of the study, autosyringe (outpatient) pumps became available and were utilized for 16% of the administered courses.

All patients treated had a measurable or evaluable metastatic site and a Karnofsky performance status of > 30%. Other criteria for patient eligibility included adequate liver function (serum bilirubin level < 2 mg/dl), unless the abnormality was due to progressive tumor and adequate bone marrow function (defined as a peripheral absolute granulocyte count > 1500/mm³ and a platelet count > 100,000/mm³), and demonstrable refractoriness to previous therapy with conventional dose schedules of vinca alkaloids. Written informed consent was obtained from all patients before therapy was started. Pretreatment evaluation included cbcs and serum chemistries (SMA-100), as well as CEA determination. Other staging workup included chest roentgenogram, complete bone survey, bone and liver scans, and photographs of soft tissue lesions.

A complete remission was defined as the disappearance of all clinical evidence of active tumor for a minimum of 4 weeks, including normalization of bone roentgenograms and absence of all tumor-related symptoms. A partial remission was defined as a > 50% reduction in the tumor including recalcification of osteolytic lesions, with no simultaneous increase in any lesion or the appearance of new lesions. In the case of hepatic metastases with hepatomegaly that could only be measured undirectionally, a 30% reduction in the hepatomegaly was required. Stable disease was defined as a reduction of < 50%, an increase of < 25%, or no change in measurable tumor for at least 2 months without the appearance of any new lesions. Progressive disease was defined as an increase of 25% in the area of measurable tumor from its smallest size during therapy or the appearance of any new lesions. All patients were considered evaluable if they had received one course of therapy and had been observed for at least 3 weeks. The Gehan modification of the Wilcoxon test (15) was used to test differences between curves.

## RESULTS

Among the 34 patients entered in the study, one patient had early death (within 2 weeks of therapy) and one patient was lost to followup. Two patients were disqualified because of major protocol violation, both having received prednisone as well as a 5-day infusion of vinblastine. Among the 30 patients who were evaluable for response, there were one complete and 11 partial responders, with an overall response rate of 40%. Responses occurred at a median of 3 weeks following the start of treatment (range, 2–9 weeks). Of the 18 nonresponders, 13 had stable disease and five had progressive disease.

Table 1 shows the response rate by previous experience with conventional dose schedules of vinca alkaloids. Nine patients had received vincristine at a dose of 1.4 mg/m² iv as part of a modified Cooper's program or in combination with hexamethylmelamine and mitomycin C. The remaining three patients had progressed while receiving vinblastine at a dose of 8 mg/m² iv every 10–14 days prior to entering the study. These three patients received continuous-infusion vinblastine at doses of 1.5–1.8 mg/m²/day. A 50% overall response rate was observed among 12 patients who had previously received vinca alkaloids, compared with a 33% response rate observed among 18 patients who had had no prior experience with vinca alkaloids ($P < 0.1$) (table 1). Only eight of the 12 patients who had received a prior standard dose schedule of vinca alkaloids were felt to have resistant disease. All of these patients had evidence of progressive disease while receiving vinca alkaloid therapy; as shown in table 2, their subsequent response to the continuous 5-day infusion of vinblastine was not significantly different from that of the four patients who were not refractory to standard vincristine therapy.

Responses by individual organ sites are shown in table 3. Lymph nodes and cutaneous lesions appeared to be the most responsive. Table 4 shows the pertinent pretreatment patient characteristics by response status. No

TABLE 1.—Previous vinca alkaloid experience

| Response | No. of patients | |
|---|---|---|
| | Prior exposure | No prior exposure |
| Complete + partial | 6 (50%) | 6 (33%) |
| Stable disease | 4 (33%) | 9 (50%) |
| Progressive disease | 2 (17%) | 3 (17%) |
| Totals | 12 | 18 |

TABLE 2.—Response by prior refractoriness to standard dose schedule of vinca alkaloids

| Prior exposure | No. of patients | |
|---|---|---|
| | Total | Responding (%) |
| Known refractory disease to— | | |
| Vincristine | 5 | 2 (40) |
| Vinblastine | 3 | 2 (66) |
| No known refractory disease to— | | |
| Vincristine | 4 | 2 (50) |
| Vinblastine | 0 | 0 |

TABLE 3.—Response by site

| Site | No. of patients | |
|---|---|---|
| | Complete + partial response | Totals (%) |
| Soft tissue | 9 | 18 (50) |
| Lymph nodes | 3 | 6 (50) |
| Lung | 2 | 7 (28) |
| Pleura | 1 | 6 (16) |
| Bone | 5 | 14 (36) |
| Liver | 1 | 3 (33) |
| Bone marrow | 0 | 2 (0) |

significant difference was found in menopausal status, number of organ sites, dominant disease sites, extent of prior therapy, or performance status. The median time from initiation of therapy to therapeutic failure was 29 weeks for the responding patients, compared with 13 weeks for the nonresponders ($P < 0.05$) (fig 1).

## Toxicity

*Hematologic toxicity (table 5).*—Myelosuppression was the most frequent toxic manifestation with this 5-day continuous-infusion vinblastine therapy. Granulocytopenia (absolute granulocyte count < 1000/mm$^3$) was more evident than thrombocytopenia (platelet count < 100,000/mm$^3$) and occurred with 75% of the administered courses at 2 mg/m$^2$/day for 5 days. However, at doses of 1.4–1.8 mg/m$^2$/day for 5 days, the myelosuppression observed was moderate, transient, and noncumulative.

*Nonhematologic toxicity (table 6).*—The incidence of gastrointestinal toxicity did not appear to be dose-related and was mild in most instances. Other side effects observed were fever during infusion, hair loss, and acute transient hepatitis. Paresthesia of the arms and legs occurred in two patients, both of whom had had prior therapy with vincristine. The most significant toxic effect was fever of unknown origin (with neutropenia) occurring in 12 patients who received vinblastine at a dose of 2 mg/m$^2$/day for 5 days; all of these patients were hospitalized for iv antibiotic therapy. Catheter-related toxic effects were infrequent; however, due to poor catheter placement, extravasation of the drug occurred in one patient resulting in erythema and swelling of surrounding soft tissue and skin which resolved within 1 week. There was no evidence of tissue necrosis, but the same patient subsequently developed axillary venous occlusion requiring removal of the venous catheter and discontinuation of infusion therapy. Another patient had cellulitis at the site of catheter insertion.

## DISCUSSION

Since its introduction in the early 1960s, there have been conflicting reports regarding the effectiveness of vinblastine used as a single agent in the treatment of advanced metastatic breast cancer (9–13). A review of the literature suggests that weekly iv injections at 0.1–0.2 mg/kg of body weight can result in an overall response rate of 20%. Although all of these studies are not strictly comparable, our observation of an overall response rate of 40% in patients with metastatic breast cancer who previously received extensive chemotherapy suggests an

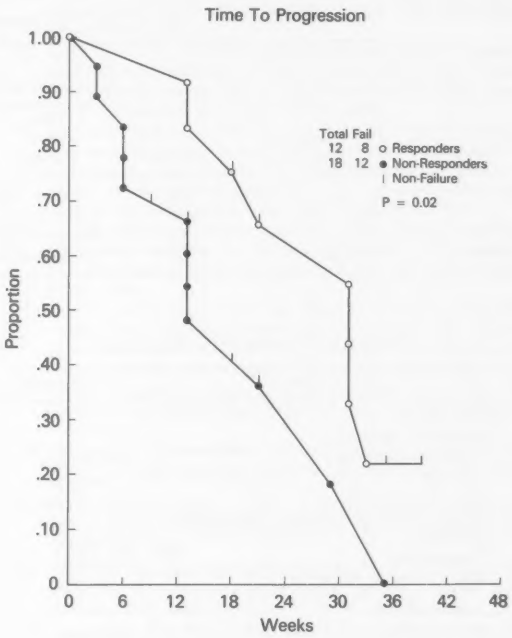



FIGURE 1.—Time from initiation of therapy to therapeutic failure.

TABLE 4.—Patient characteristics

| Characteristic | Responders | Nonresponders |
|---|---|---|
| Total No. of patients | 12 | 18 |
| Median age in yrs (range) | 53 (30–68) | 52 (34–76) |
| Median disease-free interval in days (range) | 20 (0–246) | 22 (6–360) |
| No. of patients with— | | |
| Menopausal status | | |
| Premenopausal | 2 | 4 |
| Postmenopausal | 9 | 13 |
| Unknown | 1 | 1 |
| No. of organ sites involved | | |
| 1–2 | 9 | 13 |
| 3–4 | 3 | 5 |
| Dominant disease sites | | |
| Soft tissue | 5 | 4 |
| Bone | 4 | 4 |
| Visceral | 3 | 10 |
| Median No. of prior treatment regimens (range) | 4 (1–8) | 3 (2–6) |
| Median duration of prior therapy in mos (range) | 27 (13–66) | 19 (13–48) |
| Median Karnofsky performance status (range) | 80% (50%–100%) | 80% (50%–100%) |

TABLE 5.—Hematologic toxic effects

| Dose ($mg/m^2$) | No. of courses | Lowest counts (cells × $10^3/mm^3$) | | % of courses with granulocyte count < $1000/mm^3$ | Lowest platelet count (cells × $10^3/mm^3$) (range) |
|---|---|---|---|---|---|
| | | Total wbcs (range) | Absolute granulocytes (range) | | |
| 2 | 32 | 1.7 (0.3–3.9) | 0.6 (0–1.7) | 75 | 166 (49–398) |
| 1.8 | 23 | 2.4 (1.2–4.5) | 1.2 (0–1–3.2) | 30 | 215 (40–320) |
| 1.6 | 21 | 2.2 (1.0–5.4) | 1.0 (0.2–4.2) | 34 | 210 (68–290) |
| 1.4–1.5 | 20 | 2.8 (1.3–6–1) | 1.8 (1.1–4–8) | 8 | 200 (98–315) |

TABLE 6.—Nonhematologic toxic effects

| Toxic effect | Distribution |
|---|---|
| Drug-related | |
| Nausea + vomiting | |
| Mild to moderate | 14% courses |
| Severe | 0 courses |
| Fever during infusion | 8% courses |
| Fever of unknown origin + neutropenia | 12% courses |
| Hair loss | 8 patients |
| Paresthesia | 2 patients |
| Acute transient hepatitis | 1 patient |
| Catheter-related | |
| Extravasation | 1 patient |
| Venous occlusion | 1 patient |
| Cellulitis | 1 patient |

improved therapeutic index when vinblastine is given as a prolonged constant infusion, compared with the standard dose schedule.

Mathé et al (16) and Bayssas et al (17) have reported an improved therapeutic index when vindesine is administered as a continuous 24–48-hour infusion in patients with acute lymphoblastic leukemia. More recently, in a prospective randomized study we have observed a significantly higher response rate when vindesine was given as a continuous 5-day infusion compared with the standard dose schedule (18). Similarly, in this study two of three patients who demonstrated refractoriness to intermittent doses of vinblastine of 8 $mg/m^2$ iv at 10–14-day intervals subsequently experienced a partial response to the 5-day continuous-infusion program. These results

suggest a possible augmentation of the therapeutic index of vinblastine administered by the continuous-infusion schedule. They lend support to the theoretic advantage of this schedule over the intermittent iv bolus administration. Based on initial clinical pharmacology studies, Nelson et al (14) suggested that the biweekly schedule for vinca alkaloids was superior to the weekly schedule. Additionally, we have recently demonstrated that the therapeutic efficacy of vinblastine was enhanced by sustaining a plasma drug concentration of > 1–23 μM (1 μg/ml) in the patient (19). Furthermore, the cytokinetic basis for the superiority of a continuous-infusion schedule of vinblastine was provided by animal experiments with transplantable mouse tumors, which showed that daily administrations produced antitumor effects superior to those of single large doses (20,21). Moreover, the theoretic considerations of Valeriote et al (22) suggest that the optimal effects of vinblastine require significant lengths of exposure to critical drug concentrations. Thus, the maintenance of a prolonged, constant serum concentration of vinblastine would expose more cancer cells to the drug during a susceptible phase of their cycle and might account, in part, for the higher response rate observed with this 5-day continuous-infusion program, including responses in patients whose disease was resistant to a conventional dose schedule of vinca alkaloids.

Although many patients treated in this program were hospitalized for their 5-day continuous infusion, the recent introduction of outpatient infusion pumps may make this regimen suitable for outpatients. Furthermore, the degree of antitumor activity observed with vinblastine when administered as a continuous infusion makes this a worthwhile regimen in combination with other active chemotherapeutic agents in the treatment of advanced breast cancer. Further investigations, especially in patients who had demonstrated prior refractoriness to the standard dose schedule of vinblastine, are warranted to better define the apparent increase in the therapeutic index of vinblastine when administered as a continuous infusion.

## REFERENCES


1. Canellos GP, DeVita VT, Gold GL, et al. Combination chemotherapy for advanced breast cancer: response and effect on survival. Ann Intern Med 84:389–392, 1976.
2. Haskell CM, Sparks FC, Graze PR, et al. Systemic therapy for metastatic breast cancer. Ann Intern Med 86:68–80, 1977.
3. Carbone PP, Bauer M, Band P, et al. Chemotherapy of disseminated breast cancer: current status and prospects. Cancer 39:2916–2922, 1977.
4. Young R, Lippman M, DeVita V, et al. Perspectives in the treatment of breast cancer. Ann Intern Med 86:784–798, 1977.
5. Gutterman JU, Cardenas JO, Blumenschein GR, et al. Chemoimmunotherapy of advanced breast cancer: prolongation of remission and survival with BCG. Br Med J 2(6046):1222–1225, 1977.
6. Hortobagyi GN, Yap HY, Blumenschein GR, et al. Response of disseminated breast cancer to combined modality treatment with chemotherapy and levamisole with or without bacillus Calmette-Guérin. Cancer Treat Rep 62:1685–1692, 1978.
7. Hortobagyi GN, Gutterman JU, Blumenschein GR, et al. Combination chemoimmunotherapy of metastatic breast cancer with 5-fluorouracil, Adriamycin, cyclophosphamide, and BCG. Cancer 43:1225–1233, 1979.
8. Buzdar AU, Legha SS, Tashima CK, et al. Ifosfamide versus cyclophosphamide in combination drug therapy for metastatic breast cancer. Cancer Treat Rep 63:115–120, 1979.
9. Johnson B, and Novales E. The use of vinblastine sulfate (Velban) in advanced malignancies of the female reproductive tract. Proc Am Assoc Cancer Res 5:32, 1964.
10. Bleehen N, and Jelliffe A. Vinblastine sulfate in the treatment of malignant disease. Br J Cancer 19:268–273, 1963.
11. Goldenberg I. Vinblastine sulfate (VLB) therapy of women with advanced breast cancer. Cancer Chemother Rep 29:111–113, 1963.
12. Wright TL, Hurley J, Korst DR, et al. Vinblastine in neoplastic disease: Midwest Cooperative Chemotherapy Group. Cancer Res 23:169–179, 1963.
13. Smart CR, Rochlin DB, Nahum AM, et al. Clinical experience with vinblastine sulfate (NSC-49842) in squamous cell carcinoma and other malignancies. Cancer Chemother Rep 34:31–45, 1964.
14. Nelson RL, Dyke RW, and Root MA. Comparative pharmacokinetics of the vinca alkaloids in man. Clin Pharmacol Ther 21:212, 1977.
15. Lu K, Yap HY, Watts S, et al. Comparative clinical pharmacology of vinblastine (VLB) in patients with advanced breast cancer: single versus continuous infusion. Proc Am Assoc Cancer Res and ASCO 20:371, 1979.
16. Mathé G, Misset JL, De Vassal F, et al. Phase II clinical trial with vindesine for remission induction in acute leukemia, blastic crisis of chronic myeloid leukemia, lymphosarcoma, and Hodgkin's disease: absence of cross-resistance with vincristine. Cancer Treat Rep 62:805–809, 1978.
17. Bayssas M, Gouvei G, De Vassal F, et al. Phase II clinical trials with vindesine for remission induction in leukemias and lymphomas. Proc Am Assoc Cancer Res and ASCO 20:48, 1979.
18. Yap H-Y, Blumenschein GR, Hortobagyi GN, et al. A randomized study of continuous 5-day infusion (CI) versus intermittent bolus injection (II) of vindesine in breast cancer. (Abstr 390) *In* Proceedings of the 11th International Congress of Chemotherapy—19th Interscience Conference on Antimicrobial Agents and Chemotherapy, Boston, MA, Oct 1–5, 1979.
19. Gehan EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203–223, 1965.
20. Creasey WA. Tumor-inhibitory effects of combinations of vinca alkaloids with actinomycin D. Biochem Pharmacol 15:367–375, 1966.
21. Johnson IS, Armstrong JG, Gorman M, et al. The vinca alkaloids: a new class of oncolytic agents. Cancer Res 23:1390–1427, 1963.
22. Valeriote FA, Bruce WR, and Meeker BE. A model for the action of vinblastine in vivo. Biophys J 6:145–152, 1966.

# Phase I Clinical Study of *N*-(Phosphonacetyl)-L-aspartic Acid (PALA) [1,2,3]

**Manuel Valdivieso,* E. Colleen Moore, Andrew M. Burgess, Jose R. Marti, Joseph Russ, William Plunkett, Ti Li Loo, Gerald P. Bodey, and Emil J Freireich [4,5]**

## SUMMARY

*N*-(Phosphonacetyl)-L-aspartic acid (PALA), a potent inhibitor of aspartate carbamoyltransferase (ACTase), was administered to 75 adult cancer patients to evaluate its toxicity, efficacy, and possible mechanism of action. PALA was administered at 2-week intervals as either a single or a multiple iv dose schedule. Fifty-eight patients received the single schedule at doses ranging between 0.8 and 15 g/m$^2$. Seventeen patients received the multiple-dose schedule given as 1 g/m$^2$/day for 5, 8, and 10 days respectively. Dose-limiting toxic effects for both treatment schedules were skin rash and mucositis at total doses of $> 5$ g/m$^2$. These effects appeared slightly more common with the multiple-dose schedule. Myelosuppression was negligible. The in vivo effects of PALA on tumors were studied in 18 patients with biopsiable disease. Tumor ACTase activities were lower than those of the Lewis lung carcinoma, the most sensitive animal tumor to PALA. Therapy with PALA was followed by $> 80\%$ inhibition of tumor ACTase activities and by approximately 40% inhibition of pyrimidine nucleotides. The effects of PALA on these parameters appeared to be directly related to the concentration of PALA within the tumor and inversely related to the time elapsing between the treatment and the tumor sampling. There was a partial remission in one patient with malignant melanoma and three less than partial remissions in one patient each with carcinoma of the colon and ovary and neurofibrosarcoma. Four additional patients had mixed responses and included one patient with malignant melanoma, two with colorectal cancer, and one with head and neck tumor. Antitumor activity did not seem to be schedule-dependent. We recommend that phase II studies of PALA be conducted at iv doses of 6 g/m$^2$ (single-dose schedule) or 1 g/m$^2$/day $\times$ 5 (multiple-dose schedule), repeated every 2 weeks.

[Cancer Treat Rep 64:285–292, 1980]

*N*-Phosphonacetyl-L-aspartic acid (PALA) is a transition-state inhibitor of aspartate carbamoyltransferase (ACTase) and a potent plus specific inhibitor of the biosynthesis de novo of pyrimidine nucleotides (1). Its chemical structure is shown in figure 1.

PALA was brought to the clinic with a high degree of expectation because of its broad-spectrum antitumor activity in experimental solid tumor systems, mainly the Lewis lung carcinoma and the B16 melanoma (2,3). In the Lewis lung carcinoma, PALA was highly effective

[1]Received Sept 4, 1979; accepted Oct 31, 1979.

[2]Supported in part by contract N01-CM-97277 (Division of Cancer Treatment) and Public Health Service grant CA-05831 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Presented in part at the 17th Meeting of the American Association for Cancer Research, New Orleans, LA, May 16–19, 1979.

[4]Departments of Developmental Therapeutics (*M. Valdivieso, A. M. Burgess, W. Plunkett, T. L. Loo, G. P. Bodey,* and *E. J Freireich*), Biochemistry (*E. C. Moore*), and General Surgery (*J. R. Marti* and *J. Russ*), The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston.

[5]We wish to express our indebtedness to Wanda Buffaloe and Jackie Friedman for their assistance in conducting this study and to Dr. Brian F. Issell for assistance in planning this study.

**Reprint requests to:* Manuel Valdivieso, MD, Department of Developmental Therapeutics, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, 6723 Bertner Ave, Houston, TX 77030.

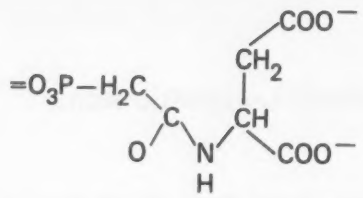


FIGURE 1.—Chemical structure of PALA.

and even curative when given at the early stages of tumor implantation. The antitumor activity of PALA was not schedule-dependent. Contrary to most agents selected for use in man, PALA has minimal or no activity in the rodent L1210 and P388 leukemias. However, collateral sensitivity of cytosine arabinoside-resistant P388 leukemia cells to PALA has recently been observed (4).

The experimental antitumor activity of PALA has been reported to be inversely related to tumor ACTase activities (3). Thus, in the Lewis lung carcinoma, for example, where PALA is highly effective and even curative, the ACTase activities were 19 ± 1 nmol/mg of protein/minute as opposed to activities of 55 ± 8 and 52 ± 8 nmols/mg of protein/minute in the rodent L1210 and P388 leukemias, where PALA has only marginal efficacy. These observations suggested that the antitumor activity of PALA could perhaps be predicted by measuring the ACTase activities of the tumors.

In animals, clinical manifestations of drug toxicity were for the most part related to gastrointestinal disturbances or central nervous system effects.[6] Among the gastrointestinal side effects, dogs and monkeys exhibited different degrees of anorexia, adipsia, weight loss, diarrhea, gastrointestinal bleeding, and dose-related changes in mucous membranes ranging from inflammation to tissue necrosis or even villous atrophy. Only relatively mild alterations of clinical chemistry parameters were observed. In the central nervous system, different degrees of hypo- or hyper-stimulation were seen, and ranged from sedation, weakness, lethargy and disorientation to unconsciousness, hypothermia, and convulsions leading to death. Of interest, reproducible hematologic toxicity was not observed. The toxic manifestations of PALA were less frequent with single doses and appeared to be more common in male mice.

Based on the above information, a phase I clinical study was conducted at our institution to determine the dose-limiting toxic effects and antitumor activity associated with the administration of PALA to man. The results of this study are presented here.

[6]Clinical brochure: PALA. Prepared by the National Cancer Institute, Bethesda, MD, Dec 1977.

## PATIENTS AND METHODS

Adult patients with advanced malignancies refractory to conventional modes of treatment who were not eligible for chemotherapy programs of higher priority were entered in this study. All patients were required to have acceptable hepatic, renal, and bone marrow functions as defined by normal serum bilirubin, normal or elevations of alkaline phosphatases and serum transaminases of < 50%, and normal BUN, serum creatinine levels, and normal peripheral blood cell counts. Metastatic disease needed to be measurable, and although biopsiable disease was preferred, it was not a prerequisite for entry in the study. Patients were aware of the investigational nature of the study and signed consent forms were obtained according to institutional policies.

PALA was obtained from the Division of Cancer Treatment, National Cancer Institute (Bethesda, MD), as 10-mg ampules of sterile aqueous solution at a concentration of 100 mg/ml. PALA was administered in a 200-ml iv solution of 5% dextrose in normal saline over 1 hour. Because of the lack of schedule dependency of PALA in experimental systems,[6,7] we investigated initially a single-dose schedule and subsequently a multiple-dose schedule. In the single-dose schedule, the initial dose of 0.8 g/m² of PALA was equivalent to one third of the low toxic dose of the most sensitive large animal, the dog. The multiple-dose schedule was started almost at the completion of the single-dose study in an effort to increase the in vivo biochemical effects of PALA in human tumors, in view of the pronounced, though transient, effects of PALA on human tumor ACTase (*see* Results section). Consequently, the multiple-dose schedule was devised to increase the duration of treatment while maintaining the daily dose of PALA at a constant level.

The initial dose of PALA in the multiple-dose schedule was 1 g/m²/day × 5 days. Courses of PALA treatment were repeated at 2-week intervals in both schedules. Independent of the schedule chosen, efforts were made to enter a minimum of three patients at each dose level of PALA and have their therapy repeated at least once at the same dose level, prior to dose escalation within the same patients. New patients were entered at the next dose escalation only after completion of a minimum observation of 2 weeks for all patients treated at the prior dose level. When a dose level was associated with clinically significant toxic effects, a minimum of six patients were entered at that dose level for a more appropriate evaluation.

Patients with accessible disease had tumor biopsy samples obtained immediately prior to and at different points after PALA treatment to determine the in vivo effects of the therapy on tumor ACTase activities and pyrimidine

[7]Johnson RK. Personal communication.

nucleotide pools. PALA in human tumors was quantitated by the ACTase inhibition method previously described for the determination of PALA in plasma (5). The measurement of tissue ACTase activities was accomplished by comparing incubation mixtures of tissue homogenates, $^{14}$C-aspartate, and carbamoyl phosphate with control incubations containing no carbamoyl phosphate. Nucleotides were separated by standard high-pressure liquid chromatography (6) and were detected by their absorbance at either 254 or 280 nm. The change of pyrimidine ribonucleotides was measured as the ratio between pyrimidine and purine ribonucleotide triphosphates and expressed as a net change of this ratio following PALA treatment. Details of these studies are being reported elsewhere.[8]

Patients were closely observed during the period of drug administration and between treatment courses for as long as they remained in the study. All side effects were recorded and analyzed. Cbcs were performed twice weekly, whereas urinalysis, serum electrolytes, and SMA-100 determinations were done once weekly. In selected patients, determinations of serum calcium and magnesium were conducted immediately before and after drug administration. Appropriate radiologic and radioisotopic examinations were usually performed every two to three treatment courses.

Tumor measurements were recorded before each treatment course and patients' responses to PALA were assessed according to usual criteria. Patients were considered to have achieved a complete response upon the disappearance of all symptoms and signs of cancer. Patients were considered to have achieved a partial remission if there was a $\geq 50\%$ reduction in the sum of the products of the two largest diameters of all measurable lesions. Patients were considered to have achieved stable disease when there was no apparent progression of a tumor which was rapidly growing prior to therapy for at least 8 weeks, or when there was tumor response which was less than a partial remission or a minimal progression of $< 25\%$. Patients were considered to have progressed when there was an unequivocal increase by $> 25\%$ in the size of any measurable lesion or the development of new lesions (this included patients who might have been responding in other metastatic sites [mixed response]).

## RESULTS

Forty male and 35 female adult patients with advanced malignant solid tumors with a median age of 55 years (range, 22–78) were entered in the study (table 1). Most patients had malignant melanoma and adenocarcinomas of the colon and lung. All patients had received extensive prior chemotherapy and had had evidence of disease progression with such therapy prior to receiving PALA.

Fifty-eight patients received 179 courses of PALA in the single-dose schedule. The number of patients receiving this schedule who were treated at each of the dose levels of PALA is shown in table 2. The dose of PALA ranged between 0.8 and 15 g/m$^2$, although most courses were given at 3.5, 5, 6, and 7.2 g/m$^2$. Seventeen patients received 34 courses of PALA (table 3) in the multiple-dose schedule. The daily PALA dose was maintained constant at 1 g/m$^2$ and it was given over 5, 8, and 10 days respectively. However, most patients received therapy over 5 and 8 days.

TABLE 1.—Patient characteristics

| | No. of patients |
|---|---|
| Diagnoses | |
| Malignant melanoma | 17 |
| Adenocarcinoma, colon | 18 |
| Adenocarcinoma, lung | 10 |
| Adenocarcinoma, breast | 6 |
| Adenocarcinoma, unknown primary | 2 |
| Adenocarcinoma, kidney, ovary, pancreas, stomach | 4 |
| Squamous cell carcinoma, lung | 6 |
| Squamous cell carcinoma, head and neck, esophagus | 2 |
| Sarcoma* | 4 |
| Multiple myeloma | 2 |
| Oat cell lung | 2 |
| Hepatoma, lymphoma† | 2 |
| Total | 75 |

*1 patient each with neurofibrosarcoma, synovial cell sarcoma, fibrosarcoma, and fibrous histiocytoma.
†Diffuse histiocytic lymphoma.

TABLE 2.—No. of patients entered at each dose level receiving PALA as a single-dose schedule

| Dose (g/m$^2$ q2wks) | Initially* | Subsequently* | No. of courses |
|---|---|---|---|
| 0.8 | 5 | 5 | 10 |
| 1.6 | 6 | 10 | 16 |
| 2.4 | 5 | 7 | 12 |
| 3.6 | 5 | 11 | 32 |
| 5 | 20 | 6 | 51 |
| 6 | 7 | 3 | 17 |
| 7.2 | 5 | 12 | 24 |
| 8.4–15 | — | 7 | 17 |

*No. of patients who at each dose level of PALA received that dose as their 1st treatment course (initially) or at any point thereafter during their PALA therapy and regardless of their initial dose (subsequently).

[8]Moore EC, Valdivieso M, Plunkett W, et al. Levels of drug in tissues, ACTase activity and pyrimidine nucleotide concentrations in biopsy specimens from patients treated with *N*-(phosphonacetyl)-L-aspartate (PALA). In preparation.

TABLE 3.—No. of patients entered at each dose level receiving PALA as a multiple-dose schedule

| Daily dose ($g/m^2$ q2wks) | Initially* | Subsequently* | No. of courses |
|---|---|---|---|
| 1 × 5 days | 8 | 1 | 18 |
| 1 × 8 days | 6 | 6 | 14 |
| 1 × 10 days | 1 | 1 | 2 |

*No. of patients who at each dose level of PALA received that dose as their 1st treatment course (initially) or at any point thereafter during their PALA therapy and regardless of their initial dose (subsequently).

The biochemical effects of PALA were studied in 18 patients who had tumor biopsies before and after PALA treatment. Those patients included eight with malignant melanoma, two each with squamous cell carcinoma and adenocarcinoma of the lung, and one each with squamous cell carcinoma of the esophagus, adenocarcinoma of the breast, neurofibrosarcoma, multiple myeloma, renal cell carcinoma, and histiocytic lymphoma. The pretreatment tumor ACTase activities in patients were substantially lower than those we observed in the Lewis lung carcinoma, and without significant differences between tumor types. The mean pretreatment ACTase levels in biopsied tissues were 0.133 (range, 0.004–0.337) compared to 0.500 nmol aspartate converted/mg of protein/minute by Lewis lung carcinoma. Following treatment with PALA, the mean percent ACTase inhibition was 82% (range, 25%–100%) and the mean percent decrease in tissue pyrimidine ribonucleotides was 39% (range, 0–72%). In 16 patients, these effects were investigated further to include their possible relationship with PALA doses, PALA tumor concentrations, and time of the biopsy in relation to the therapy (table 4). The effects of PALA on tumor ACTase and pyrimidine nucleotides appeared to be directly related to the concentrations of PALA within the tumor and inversely related to the time elapsing between the treatment and the tumor sampling. The effects of the dose of PALA on these parameters as well as that of the schedule of administration are not yet well-understood.

Sixty of the 75 patients entered were evaluable for response. Reasons for unevaluability included death from tumor within 1–2 weeks from initiation of therapy, seven patients; lost to followup, four patients; and inadequate therapy (less than one full treatment course), four patients. PALA demonstrated antitumor activity in 17 patients, including one patient achieving a partial remission, three achieving less than partial remissions, nine achieving disease stabilizations, and four achieving mixed responses (table 5). The partial remission occurred in a patient with malignant melanoma and a 3-cm metastatic nodule in the breast. Tumor regression of > 50% was achieved with 7.2-$g/m^2$ PALA doses given as single bolus injections. Response lasted 4 weeks.

Less than partial remissions were observed in three patients: one with pulmonary and hepatic metastases due to colorectal cancer, one with massive ascites due to ovarian carcinoma, and one with multiple skin lesions due to neurofibrosarcoma. These three patients responded with multiple PALA doses of 1 $g/m^2$/day × 8 days (colorectal cancer) and single bolus injections of 2.4 $g/m^2$/day (ovarian cancer) and 5 $g/m^2$/day (neurofibrosarcoma). Of the four patients with mixed responses, two received single bolus injections of 7.2 $g/m^2$ (colorectal cancer and squamous cell carcinoma of the head and neck) and two received a daily dose of 1 $g/m^2$ × 5 (malignant melanoma) and 1 $g/m^2$ × 8 (colorectal cancer).

It was only possible to study the in vivo biochemical effects of PALA in the patient with neurofibrosarcoma (less than a partial remission) and in a patient with malignant melanoma (mixed response). In the patient with neurofibrosarcoma, the pretreatment tumor ACTase was 0.084 nmol aspartate converted/mg of protein/minute. The tumor PALA concentration 20 hours after therapy was 8.3 ng/g; the percent ACTase inhibition after therapy was 79% and the percent decrease in tissue pyrimidine ribonucleotides following therapy was 72%. In the patient with malignant melanoma, the tumor ACTase before treatment was 0.295 nmol aspartate converted/mg of protein/minute, the tumor PALA concentration 96 hours after completion of therapy was 2.7 ng/g, and the percent ACTase inhibition at that point was 72%. Studies of tumor pyrimidine nucleotides could not be conducted in this patient. The in vivo effects of PALA in these two patients were not significantly different from those observed in the remaining 16 patients who had tumor biopsy studies, of whom 14 had tumor progression (including three who

TABLE 4.—Effects of PALA on human tumors*

| Dose range ($g/m^2$) | No. of samples | Interval between biopsies in hrs (range) | Mean PALA concentration (μg/g) | Mean % ACTase inhibition | Mean % decrease in UTP and CTP† |
|---|---|---|---|---|---|
| 5–8 | 7 | 1.5–6 | 46 | 95 | 58 |
| 3.6–8 | 7 | 16–48 | 6.7 | 77 | 43 |
| 3.6–6 | 2 | 96–145 | 1.8 | 49 | Not Done |

*Summarized from footnote 8.
†UTP = uridine 5′-triphosphate and CTP = cytidine 5′-triphosphate.

TABLE 5.—Antitumor activity of PALA

| | No. of patients | No. of responses | | |
|---|---|---|---|---|
| | | Partial remission | Stable disease | Disease progression |
| Malignant melanoma | 12 | 1 | 3 | 8 (1 mixed) |
| Adenocarcinoma, colon | 16 | | 5 (1 < PR)* | 11 (2 mixed) |
| Adenocarcinoma, lung | 8 | | 1 | 7 |
| Adenocarcinoma, breast | 6 | | | 6 |
| Adenocarcinoma, unknown primary | 2 | | | 2 |
| Adenocarcinoma, kidney | 1 | | | 1 |
| Adenocarcinoma, ovary | 1 | | 1 (< PR)* | |
| Squamous cell carcinoma, lung | 4 | | | 4 |
| Squamous cell carcinoma, head and neck | 1 | | | 1 (mixed) |
| Squamous cell carcinoma, esophagus | 1 | | | 1 |
| Sarcoma | 4 | | 2 (1 < PR)* | 2 |
| Histiocytic lymphoma | 1 | | | 1 |
| Small cell carcinoma, lung | 2 | | | 2 |
| Hepatoma | 1 | | | 1 |
| Totals | 60 | 1 | 12 (3 < PR)* | 47 (4 mixed) |

*< PR = less than a partial remission.

died during the first course of therapy) and two had stable disease.

The toxic manifestations observed in this study in relation to the treatment schedule of PALA are shown in tables 6 and 7. In both schedules, the toxic effects were similar, although they tended to be more common at lower doses in patients receiving the multiple-dose schedule. Myelosuppression was unusual and not dose-related. The most common manifestations of drug toxicity were skin rash and mucositis which became dose-limiting at total doses of $> 5$ g/m$^2$. The skin lesions consisted of a mild and transient maculopapular rash appearing approximately 3 days after PALA therapy and involving the face, neck, and/or arms and legs. In severe cases, the rash became generalized and patients developed painful bullous-type lesions which subsequently became exfoliative and healed in 1–2 weeks, leaving areas of hyperpigmentation. All episodes of mucositis involved the oral mucosa and in one half of the cases it involved the rectum, vagina, bladder, and conjunctivas. Especially disturbing was the development of superficial and reversible episodes of keratitis in two patients who received the 8-day multiple-dose schedule. Nine of the ten episodes of severe skin rash, and all six episodes of severe mucositis occurred at total doses of $\geq 5$ g/m$^2$. The frequency and severity of these side effects did not appear to be related to gross biochemical abnormalities of hepatic or renal functions or to the in vivo effects of PALA on tumor ACTase activities and pyrimidine nucleotide pools.

During drug administration we noticed, among other toxic effects, brief episodes of paresthesias in the perioral area as well as in the arms and legs. These toxic effects lasted $<$ 3–5 minutes and occurred at total doses of $\geq 2.4$ g/m$^2$. They appeared to be related to the rapidity of drug administration and not to immediate changes in serum calcium and magnesium which were studied in ten patients.

## DISCUSSION

This phase I clinical study has identified acceptable dose-limiting toxic effects and in vivo biochemical and antitumor effects of PALA. As in animal studies, the administration of PALA to man is devoid of myelosuppressive properties, although it produced significant skin and gastrointestinal toxic effects. Central nervous system toxicity, which is dose-limiting in animals, was not observed in this study.

The toxic manifestations of PALA in the skin and mucosa were usually seen at doses of $\geq 3.6$ g/m$^2$/day $\times$ 1, and became prohibitive at doses of $\geq 7.2$ g/m$^2$ in the single-dose schedule and 1 g/m$^2$/day $\times$ 8 in the multiple-dose schedule, repeated every 2 weeks. The skin lesions usually consisted of a mild and transient maculopapular rash appearing on the face, neck, and/or arms and legs. In severe cases, patients developed painful and generalized bullous-type lesions which subsequently became exfoliative and healed, leaving areas of hyperpigmentation. The mucositis was usually oral, mild, and appeared concomitantly with the skin lesions; when it became severe, it prevented patients from eating and drinking and extended to involve the vagina, bladder, and rectal mucosa. In two instances, a transient, though troublesome, keratitis developed. During drug administration we observed, among other toxic effects, paresthesias around the mouth and on the arms and legs at total doses of $\geq 5$ g/m$^2$. The development of these toxic effects was not clearly explained on the basis of obvious biochemical ab-

TABLE 6.—Toxic effects: single-dose schedule of PALA

| | Dose (g/m²/day × 1 q2wks) | | | | | |
|---|---|---|---|---|---|---|
| | 0.8, 1.6, 2.4 | 3.6 | 5 | 6 | 7.2 | 8–15 |
| No. of evaluable courses | 37 | 27 | 39 | 15 | 24 | 10 |
| Courses with toxicity | 30% | 59% | 54% | 80% | 83% | 70% |
| No. of patients with— | | | | | | |
| Toxic effects | | | | | | |
| Myelosuppression* | 2 | — | 2 | — | — | — |
| Skin rash | 1 (3%) | 12 (44%) | 6 (15%) | 6 (40%) | 10 (42%) | 3 (30%) |
| Severe | — | 1 | 2 | 1 | 5 | — |
| Mucositis | — | 3 (11%) | 3 (8%) | 4 (27%) | 7 (33%) | 6 (60%) |
| Severe | — | — | — | 1 | 3 | — |
| Paresthesias | — | — | 9 (23%) | 6 (40%) | 5 (21%) | — |
| Nausea and vomiting | 1 | 4 | 3 | 4 | 3 | — |
| Diarrhea | — | 3 | 1 | 1 | 3 | — |
| Phlebitis | — | — | 1 | — | 2 | — |
| Fever | 2 | — | — | — | — | — |
| Anorexia | 1 | — | — | — | 1 | — |
| Dizziness | — | — | — | 1 | — | — |
| Upper gastrointestinal bleeding† | — | — | 1 | — | — | — |
| Weakness | — | — | — | — | 1 | 1 |
| Chills | — | — | — | 2 | — | — |

*Leukocyte count ≤ 3000/mm³, neutrophil count ≤ 1000/mm³, and platelet count ≤ 100,000/mm³.
†Occurred in a patient whose platelet count dropped from 41,000 to 23,000/mm³.

normalities of hepatic and renal functions or of changes in serum electrolytes or tumor ACTase activities. The schedule of drug administration appeared to influence the manifestations of drug toxicity, meaning, mainly, that they tended to occur more frequently at lower doses in the multiple-dose schedule. Studies at other institutions would also suggest that PALA is probably more toxic to the gastrointestinal tract when given by prolonged iv infusion (7,8).

Tumor ACTase activities as well as concentrations of PALA and pyrimidine ribonucleotides in tumor samples were determined in 18 patients entering this study to evaluate the possible mechanism of action of PALA and also to investigate in man the applicability of the preclinical observations of Johnson et al (3) that the antitumor effects of PALA were inversely related to the pretreatment tumor activities of ACTase. In comparison with the ACTase activities in the Lewis lung carcinoma, where PALA is most effective and even curative, the pretreatment median and mean values in man were even lower. In addition, most patients had > 80% inhibition of tumor ACTase after PALA treatment. The degree of tumor ACTase inhibition paralleled the concentration of PALA within the tumor, being greatest in the first 6 hours after therapy and progressively lower thereafter. Similarly, PALA treatment produced a significant reduction (58%) in tumor tissue ribonucleotides during the first 6 hours of therapy. The changes in tumor ribonucleotides, however, did not always correlate with ACTase changes, so that at the present time, a direct cause and effect relationship cannot be deduced from these studies. More details on these observations are being reported separately.[9]

The antitumor activity of PALA in this study was modest and did not appear to be related to the schedule of drug administration. A partial remission was observed in a patient with malignant melanoma and soft tissue disease, and three less than partial remissions were observed in one patient each with carcinomas of the colon and ovary and neurofibrosarcoma. Four additional patients, including one with malignant melanoma, two with colorectal cancer, and one with head and neck tumor, had mixed responses. The limited efficacy of PALA in this study cannot be easily explained based on the available data. The experimental activity of PALA in tumors containing low ACTase levels would have predicted a greater

[9] *See* footnote 8.

TABLE 7.—Toxic effects: multiple-dose schedule of PALA

| | Dose q2wks | |
|---|---|---|
| | 1 g/m²/day × 5 | 1 g/m²/day × 8 |
| No. of evaluable courses | 14 | 13 |
| Courses with toxicity | 69% | 100% |
| No. of patients with— Toxic effects | | |
| Myelosuppression* | — | — |
| Skin rash | 6 (38%) | 9 (69%) |
| Severe | 1 | — |
| Mucositis | 5 (31%) | 7 (54%) |
| Severe | 1 | 1 |
| Paresthesias | 2 | — |
| Nausea and vomiting | 3 | 3 |
| Diarrhea | 2 | 1 |
| Weakness | — | 1 |
| Bleeding | 2 | — |
| Anorexia | — | 2 |
| Anxiety | — | 1 |
| Keratitis | — | 2 |

*Leukocyte count ≤ 3000/mm³, neutrophil count ≤ 1000/mm³, and platelet count ≤ 100,000/mm³.

antitumor effect in man in view of their even lower levels of tumor ACTase than those of the most sensitive animal tumor to PALA, the Lewis lung carcinoma. Greater antitumor activity would have also been expected in view of the significant degree of tumor ACTase inhibition and to the lesser degree of pyrimidine nucleotides following PALA treatment. Since PALA is not always effective against animal tumors with low ACTase activities (3), the possibility exists that a mechanism other than ACTase inhibition is involved in this reaction. The so-called salvage pathways, synthesis of nucleotides from exogenous or recycled uridine, cytidine, or uracil, represent a possible explanation. In fact, tumors highly efficient in utilizing circulating uridine would be less sensitive to PALA, since this nucleoside can effectively reverse the effects of PALA in vitro and in vivo (9,10).

The administration of PALA to man appears to be well-tolerated. Myelosuppression is negligible and toxic effects to the skin and mucosas are dose-limiting at higher doses. PALA possesses significant biochemical effects in human tumors as demonstrated by substantial though transient inhibitions of tumor ACTase and, to a lesser degree, of pyrimidine nucleotides. In spite of this, the observed antitumor activity of PALA is thus far modest, suggesting that blocking the de novo pyrimidine biosynthesis under the conditions of this study may not be sufficient to obtain tumor regression in man. Therefore, in addition to phase II studies, future consideration should be given to studying PALA in combination with other agents in an effort to increase the antitumor effects of these agents by providing additional inhibitory steps in the nucleotide biosynthesis of tumor cells. For initial phase II studies of PALA, we recommend iv doses of 6 g/m²/day × 1 or 1 g/m²/day × 5 repeated every 2 weeks.

## REFERENCES


1. COLLINS KD, and STARK GR. Aspartate transcarbamylase. Interaction with the transition state analogue *N*-(phosphonacetyl)-L-aspartate. J Biol Chem 246:6599–6605, 1971.
2. JOHNSON RK, INOUYE T, GOLDIN A, ET AL. Antitumor activity of *N*-(phosphonacetyl)-L-aspartic acid, a transition-state inhibitor of aspartate transcarbamylase. Cancer Res 36:2720–2725, 1976.
3. JOHNSON RK, SWYRYD EA, and STARK GR. Effects of *N*-(phosphonacetyl)-L-aspartate on murine tumors and normal tissues in vivo and in vitro and the relationship of sensitivity to rate of proliferation and level of aspartate transcarbamylase. Cancer Res 38:371–378, 1978.
4. LASTER WR, and SCHABEL FM, JR. Collateral sensitivity of P388/Ara-C and P388/5-FU to *N*-(phosphonacetyl)-L-aspartate (PALA). Proc Am Assoc Cancer Res and ASCO 20:95, 1979.
5. FRIEDMAN J, MOORE EC, HALL SW, ET AL. Enzymatic assay for the antitumor agent *N*-(phosphonacetyl)-L-aspartic acid (PALA). Cancer Treat Rep 63:85–88, 1979.
6. PLUNKETT W, ALEXANDER L, CHUBB S, ET AL. Comparison of the activity of 2'-deoxycoformycin and erythro-9-(2-hydroxy-3-nonyl)adenine in vivo. Biochem Pharmacol 28:201–206, 1979.

7. Ervin TJ, Blum RH, and Canellos GP. *N*-(Phosphonacetyl)-L-aspartate (PALA), phase I trial. Proc Am Assoc Cancer Res and ASCO 20:200, 1979.
8. Ohnuma T, Hart R, Roboz J, et al. Clinical and pharmacological studies with phosphonacetyl L-aspartate (PALA). Proc Am Assoc Cancer Res and ASCO 20:344, 1979.
9. Johnson RK. Reversal of toxicity and antitumor activity of *N*-(phosphonacetyl)-L-aspartate by uridine or carbamyl-DL-aspartate in vivo. Biochem Pharmacol 26:81–84, 1977.
10. Swyryd EA, Seaver SS, and Stark GR. *N*-(Phosphonacetyl)-L-aspartate, a potent transition state analog inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. J Biol Chem 249:6945–6950, 1974.

# Effect of Adriamycin on the Cell Kinetics of 13762 Rat Mammary Tumors and Implications for Therapy[1,2]

Paul G. Braunschweiger* and Lewis M. Schiffer[3,4]

## SUMMARY

**The unperturbed cell kinetics of 13762 mammary tumors in Fischer 344/CRBL rats were studied by fraction labeled mitoses (FLM) and in vitro methods. The $^3$H-thymidine ($^3$H-TdR) labeling index (LI) was 32.5% and the volume doubling time was approximately 96 hours 14 days after inoculation. The cell cycle, $G_1$, $G_2$, and DNA synthesis times by computer analysis of the FLM data were 11.5, 3.5, 1.3, and 5.4 hours respectively. The primer available DNA-dependent DNA polymerase labeling index (69%) was similar to the growth fraction calculated from the FLM data. Estimates of cell kinetic parameters from this tumor model, by in vitro methods, were very similar to those obtained by FLM analysis. The effect of Adriamycin (Adr) on the cell kinetics of 13762 tumors was also studied by monitoring the changes in the $^3$H-TdR LI, the primer available DNA-dependent DNA polymerase labeling index, and the flux of cells from S to $G_2$ phase by the $^3$H-TdR and $^{14}$C-TdR double-labeling method. The results indicated that cell proliferation was inhibited for at least 3 days after 5 mg/kg of Adr. Recovery of cell proliferation was noted between Days 4 and 7 after treatment. Sequential chemotherapy with cyclophosphamide was most effective when administered to coincide with the proliferative recovery after Adr.**

[Cancer Treat Rep 64:293–300, 1980]

Adriamycin (Adr), an anthracycline antibiotic derived from *Streptomyces peucetius*, is a 14-hydroxyl derivative of daunorubicin (1,2). The drug is found to be associated with the nucleoprotein fraction in treated cells (3), and its principal mode of action is thought to be inhibition of DNA and RNA synthesis by intercalation with DNA (4). Although cultured cells in S phase (5–8) or at the S-$G_2$ border (9) are most sensitive to this agent, lethality may result from an irreversible $G_2$ block (8,10,11). Adr has activity against a wide range of experimental tumors (4,12) including C3H mammary tumors (13) and many human tumor types (14). Cumulative response rates of 43% and 26% have been reported in breast cancer for untreated and previously treated patients respectively (15). Adr has also been used both in the laboratory and in the clinic as one component of multiple-drug protocols. The most actively studied and perhaps most promising combinations have included cyclophosphamide (Cp) (10,16–23).

Previous studies have shown that tumor cell kinetics after perturbation with chemotherapeutic drugs (19, 20,24–28) and radiation therapy (29) can be useful in designing more effective sequencing intervals. In the 13762 rat mammary tumor (RMT), Cp administered sequentially at a kinetically defined interval (9 days) was substantially more effective than similar total drug doses given at one time (20). In the present studies, the cell

[1]Received Sept 7, 1979; revised Nov 16, 1979; accepted Dec 6, 1979.

[2]Supported by contract N01-CB-43899 (Breast Cancer Task Force) and in part by Public Health Service grant CA-10438 (Clinical Radiation Therapy Research Center) from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Cancer Research Unit, Clinical Radiation Therapy Research Center, Allegheny General Hospital, Pittsburgh, PA.

[4]We would like to thank Saundra Banks, Barbara Grauel, Tamara Hamilton, Melissa Patterson, Agnese Pollice, Kathleen Simpson, and Linda Stragand for their technical assistance, Katherine Pawlos for preparation of the manuscript, and Dr. G. Gordon Steel for computer analysis of the fraction labeled mitoses data.

*Reprint requests to: Dr. Paul G. Braunschweiger, Cancer Research Unit, Division of Radiation Oncology, Allegheny General Hospital, 320 E North Ave, Pittsburgh, PA 15212.

kinetics were studied before and from 6 hours to 7 days after treatment with 5 mg/kg of Adr. Therapy protocols based on the Adr-induced kinetic perturbations were designed and tested.

## MATERIALS AND METHODS

### Animals and Tumor Model

Fischer 344/CRBL female rats, 6–8 weeks of age, were used throughout the experiments. The 13762 RMT was originally induced with 7,19,dimethylbenz(α)anthracene (30). For our initial baseline fraction labeled mitoses (FLM) studies, tumor-bearing rats were kindly supplied by Dr. Arthur Bogden (Mason Research Institute, Worcester, MA). The tumor was subsequently maintained by subcutaneous implantation of 2–3-$mm^3$ tissue fragments, on the flank, every 2 weeks. Tumor-bearing rats were housed three to four per cage and were fed water and lab chow (Ralston Purina, Evanston, IL) ad libitum.

Histologically, the 13762 RMT is a moderately well-differentiated adenocarcinoma with acini showing papillary growth and central necrosis in older tumors. Tumors were measured with calipers in two dimensions and sizes were expressed as a mean of the two diameters (31). Duplicate measurements were made at least twice weekly, but daily measurements were usually made following treatment. Tumor volumes were calculated using the equation for spherical volume, and volumetric doubling times were estimated.

### Drug Treatments

Tumor-bearing rats were weighed and given 5 mg/kg of Adr (Adria Laboratories, Wilmington, DE) in sterile saline by ip injection 14 days after transplantation when tumors were approximately 2.7 cm in diameter. Animals were killed at various times after treatment for the cell kinetic studies. At least four animals were used at each study interval from 6 hours to 7 days after treatment. Cp (80 or 160 mg/kg), used in some studies, was also dissolved in sterile saline and was administered ip. All drugs were freshly prepared prior to use.

### Cell Kinetic Analysis

#### *Fraction Labeled Mitoses*

Forty-eight rats with 14-day-old 13762 RMTs were pulse-labeled with $^3$H-thymidine ($^3$H-TdR) (1 microcurie/g of body weight, 8 curies/mmol, Schwarz/Mann, Orangeburg, NY) and were killed at 30 minutes and 1-, 1.5-, 2-, and 2-hour intervals thereafter for the next 22 hours. At the appropriate times rats were killed by cervical dislocation, tumors were resected and fixed in formalin, and autoradiograms of Feulgen-stained paraffin sections were prepared. The FLM was determined for each of the three tumors per interval by counting 100 mitoses/animal. Cells containing > 3 grains/cell were considered labeled. Background was consistently less than 1 grain/cell and mean grain counts were always > 20 grains/cell.

### In Vitro Methods

The methods for in vitro single and double labeling of solid tumors have been previously described (32). Briefly, tumor-bearing rats were killed by cervical dislocation. The tumors were resected and minced in McCoy's medium with 10% fetal calf serum (GIBCO, Grand Island, NY). The resultant cell suspensions were filtered through a nylon gauze to remove large tissue fragments and were then either single labeled with $^3$H-TdR (15–17 curies/mmol, 2.5 microcuries/ml) for 1 hour or double labeled for the determination of the DNA synthesis time with $^3$H-TdR as above and then for an additional 30 minutes with $^{14}$C-Tdr (54 millicuries/mmol, 0.25 microcurie/ml). All labeling in vitro was at 37°C in a shaking water bath. Following the labeling interval, the viable cells were separated by Ficoll-Hypaque gradient centrifugation and slide preparations were made for autoradiography.

The primer available DNA-dependent DNA polymerase (PDP) assay is an autoradiographic method to measure the simultaneous presence of nuclear DNA polymerase and nuclear DNA primer template activity in individual cell nuclei. This technique has also been previously described (33–36). Imprints from a freshly cut non-necrotic tumor surface were made on microscope slides and air-dried. The cytoplasm was stripped by dipping in 0.125% agar solution at 40°C. This process appears to remove the cytoplasm, leaving only bare nuclei adherent to the slide. The slides were then fitted with glass incubation chambers into which was put 0.5 ml of an incubation mixture containing the deoxy-5′-triphosphates of adenosine, guanosine, and cytidine (0.2 mM), 5 mM $MgCl_2$, Ficoll (Sigma Chemical Co, St. Louis, MO), and 5 microcuries of $^3$H-thymidine triphosphate (54 curies/mmol, New England Nuclear Corp, Boston, MA) in 0.02 M tris buffer (pH 7.4). The preparations were incubated at 37°C for 45 minutes at which time the reaction was stopped on ice. The incubation chambers were then removed, the slides were rinsed, and the cells were fixed in acidic formaldehyde and rinsed prior to autoradiography.

Kodak NTB-2 liquid photographic emulsion was used throughout the experiments. The slides were dipped in emulsion at 45°C, air-dried, and then exposed for 1 and 7 days for single label and PDP slides respectively. $^3$H-TdR labeling indices (LIs) and primer available DNA-dependent DNA polymerase labeling indices (PDPIs) were determined by examining at least 500 cells. Double-label samples were assayed by double-emulsion autora-

diography as previously described (32). Autoradiograms of sections used for FLM analysis were exposed for 4 or 6 weeks.

### Calculations

Calculation of cell kinetic parameters for the in vitro data was done as previously described (31,33,35) using the formulations of Wimber and Quastler (37) for the DNA synthesis time and the formulations of Steel (38) for the cell cycle time and other cell kinetic parameters. The FLM data were subjected to computer simulation by the method of Steel and Hanes (39).

The double-label method as used in the present experiments can also be used to estimate the DNA synthesis time ($T_S$) in synchronously proliferating populations (37). After drug treatments, cell cycle distributions may be nonrandom, thus invalidating the $T_S$ calculation. The ratio of double-labeled to single-labeled cells provides an index (DLI) as to the movement of cells from S to $G_2$ phase at the time of assay (40).

## RESULTS

### Unperturbed Cell Kinetics

Cell kinetic parameters for the unperturbed 14-day 13762 RMTs determined by FLM analysis and computer simulation are shown in table 1. Cell kinetic parameters determined by our in vitro methods are also shown in table 1. FLM analysis showed median values for $T_S$, $G_1$ time ($T_{G1}$), $G_2$ time ($T_{G2}$), and cell cycle time ($T_C$) of 5.4, 3.5, 1.8, and 11.5 hours respectively. The growth fraction (GF) calculated from $LI_{observed}/LI_{theoretical}$ was approximately 69%. This rapidly growing tumor model (volumetric Td = 96 hours) shows a high degree of cell loss, with a cell loss to cell production ratio (Ø) of approximately 0.84.

TABLE 1.—Cell kinetic parameters for 13762 RMT by in vivo (FLM) and in vitro methods

| Parameter* | In vivo | In vitro |
|---|---|---|
| Td | 96 hrs | |
| MI | 2.17 ± 0.15† | 1.92 ± 0.16 |
| LI | 32.5 ± 2.6 | 31.1 ± 2.1 |
| $T_S$ | 5.4 (0.1)‡ | 5.7 ± 0.2 |
| $T_C$ | 11.5 | 11.9 ± 0.6 |
| $T_{G1}$ | 3.5 (1.0) | 4.1 ± 0.2§ |
| $T_{G2}$ | 1.8 (0.6) | 1.8 |
| $T_M$ | 0.8 | 0.4¶ |
| Tp | 14.4 | 15.5 ± 0.8 |
| GF** | 69.2 | |
| PDPI | | 69.0 ± 2.0 |
| λ | 0.865 | 0.860 |
| $K_p$ | 0.0482 | 0.0447 |
| $K_L$ | 0.0407 | 0.0375 |
| Ø | 0.844 | 0.840 |

*MI = mitotic index; $T_M = T_C - [T_S + T_{G1} = T_{G2}]$; Tp = potential doubling time; GF = $LI/[T_S/T_C]$; λ = correction for nonlinear age distribution; Ø = cell loss factor.

† Mean ± 1 SE.

‡Nos. in parentheses = coefficient variation.

§$T_{G1} = T_C - [T_S + T_{G2} + T_M]$.

¶$T_M = [T_S]\ [\lambda]\ [MI]/[LI]\ [1n2]$.

In vitro assays for LI, $T_S$, and PDPI were performed for individual tumors and the mean $T_C$ was calculated from at least five tumors. Cell kinetic parameters determined in this way were very similar to those obtained by FLM analysis. As in other tumor models, the PDPI was a good estimate of the GF for the 13762 RMT.

### Tumor Growth

Figure 1 shows the effect of Adr (5 mg/kg) and Cp (80 mg/kg) on the growth of 13762 RMTs. The mean diameter is expressed as a fraction of the treatment day size (2.7 cm).

Treatment with 5 mg/kg of Adr did little to affect a change in tumor growth. A single dose of Cp (80 mg/kg), however, resulted in a 51% decrease in tumor diameter with approximately a 29-day regrowth delay. Simultaneous administration of both agents resulted in tumor shrinkage and regrowth delay similar to that seen with Cp alone.

### Perturbed Cell Kinetics

Figure 2 shows the effect of 5 mg/kg of Adr on the $^3$H-TdR LI, the PDPI, and the DLI in 13762 RMTs. Although the $^3$H-TdR LI is reduced at 6 hours, the PDPI is still normal. Twenty-four hours after Adr was given both the $^3$H-TdR LI and PDPI were subnormal. $^3$H-TdR LIs and PDPIs remained subnormal until Day 4 when both parameters were within normal range. The DLI was only slightly affected by Adr at the intervals studied. The most significant change was a decrease of 1 hour seen at 24 hours.

### Protocol Design

To evaluate the therapeutic implication of the Adr-induced cell kinetic perturbations, subsequent treatment was initiated during intervals which from our kinetic data indicated intervals of subnormal proliferation (Days 1 and 2 after Adr) or proliferative recovery (Day 5). Although the kinetic data indicated recovery on Day 4 after Adr, Day 5 was chosen to minimize normal tissue toxicity potentials (41). Some of the results are shown in figure 3.

Cp alone resulted in about 50% tumor shrinkage and a 29-day regrowth delay. Cp given 2 days (Day 16) after Adr (Day 14) resulted in less tumor shrinkage and a shorter regrowth delay period (19 days from the time of Cp treatment). Cp given 5 days after Adr resulted in little or no overt shrinkage, but a 21-day regrowth delay

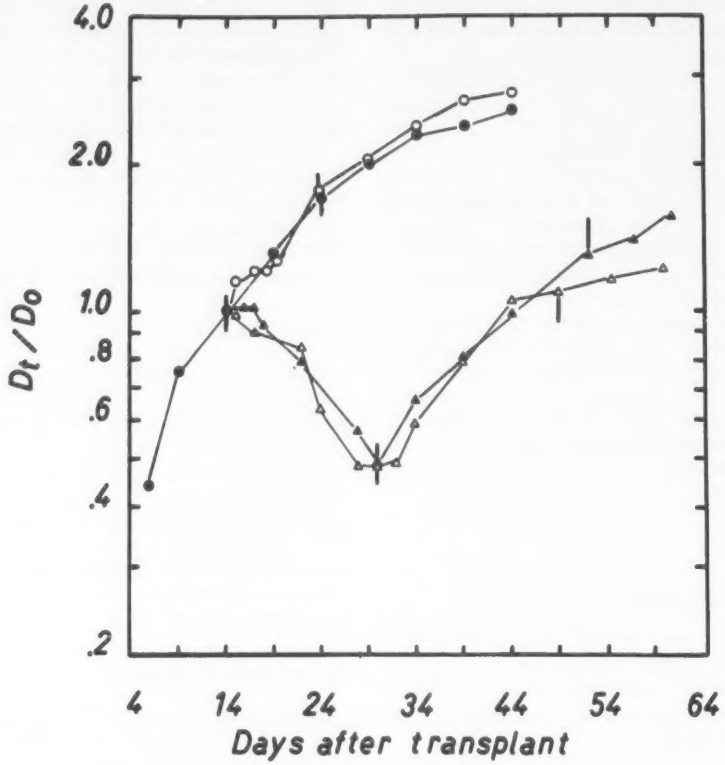



FIGURE 1.—Effect of Adr (○) and Cp (△) alone and in combination (▲) on tumor growth. The growth of untreated tumors (●) is also shown. Tumor diameter ($D_t$) is expressed as a fraction of the 14-day pretreatment ($D_o$) value. Each point = mean ± 1 SE for 4–5 tumors. Coefficients of variation were < 15%.

from the time of Cp treatment was noted. Tumor regrowth between 34 and 64 days was most rapid in tumors treated with Cp alone and was slowest in tumors treated with Cp 5 days after Adr.

Other tumor-bearing rats were treated with Adr and then either 1 or 5 days later were treated with two doses of Cp (80 mg/kg) administered at a previously defined kinetically based sequencing interval of 9 days. Although rapid tumor regression and complete remissions were observed when the first Cp dose was given 1 day after Adr, tumor regrowth was soon observed in all rats. When the two-dose Cp schedule was initiated on Day 19 to correspond to the interval of proliferative recovery, shrinkage was slower, but all animals attained a complete and long-lasting remission.

Table 2 shows a summary of protocols in which Cp was administered sequentially during intervals of subnormal proliferation and proliferative recovery after Adr. Adr alone was ineffective in producing either cures or long-term survivors. Cp given 0, 2, or 5 days after Adr resulted in increased survival time as compared to Cp alone. The mean survival time was, however, not different for the three Adr plus Cp schedules. One long-term survivor was noted in the Adr (14-day) plus Cp (19-day) study.

Two doses of Cp (Days 14 and 23) yielded 80% complete remissions but only 30% long-term survivors. When the Cp schedule was initiated 5 days after Adr, substantially better cure rates (100%) were noted. Although the Cp schedule initiated 1 day after Adr produced complete remissions, no long-term survivors were noted.

## DISCUSSION

In the present studies the cell cycle kinetics of 13762 RMTs were studied prior to and at various times after 5 mg/kg of Adr. This dose is approximately equivalent

TABLE 2.—Efficacy of cell kinetic-based chemotherapy in 13762 RMTs

| Protocol No. | Adr* (Day) | Cp* (Day) | Adr alone (mg/kg) | Cp alone (mg/kg) | MST† | Remission‡ (%) | Cure§ (%) | Total No. of animals |
|---|---|---|---|---|---|---|---|---|
| 1 | | | 0 | 0 | 32.6±4.1 | 0 | 0 | 10 |
| 2 | 14 | | 5 | | 38.0±2.0 | 0 | 0 | 8 |
| 3 | | 14 | | 80 | 50.6±3.9 | 25 | 5 | 20 |
| 4 | 14 | 14 | 5 | 80 | 65.3±6.3 | 25 | 0 | 4 |
| 5 | 14 | 16 | 5 | 80 | 73.3±4.0 | 0 | 0 | 4 |
| 6 | 14 | 19 | 5 | 80 | 60.6±11.3 | 25 | 25 | 4 |
| 7 | | 14 | | 160 | 73.4±15.2 | 0 | 0 | 4 |
| 8 | | 14,23 | | 80 | 88.9±7.8 | 80 | 30 | 23 |
| 9 | 14 | 19,28 | 5 | 80 | >130 | 100 | 100 | 4 |
| 10 | 14 | 15,24 | 5 | 80 | 87.0±8.1 | 100 | 25 | 4 |

*Day after transplantation when drug was administered.
†Mean survival time ± 1 SE.
‡% of animals showing complete remission.
§% of animals living tumor-free for > 100 days.

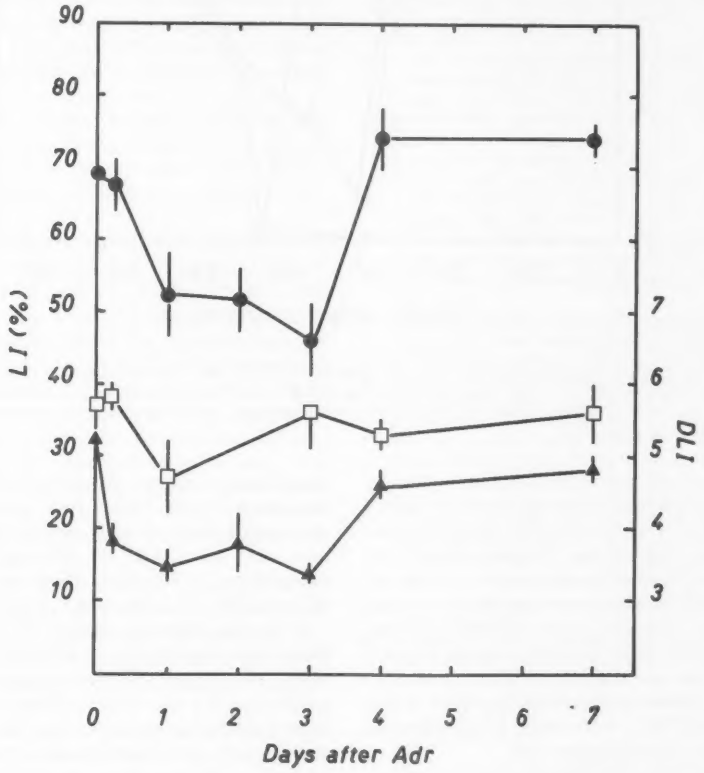



FIGURE 2.—Effect of Adr (5 mg/kg) on $^3$H-TdR LI (▲), PDPI (●), and DLI (□) at various intervals after treatment. Each point = mean ± 1 SE for 4 tumors.

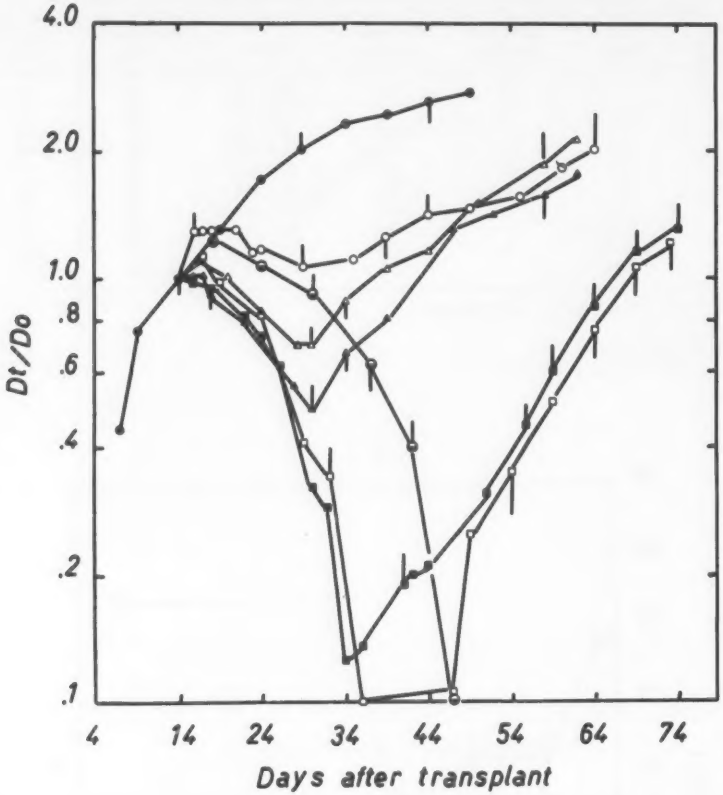



FIGURE 3.—Effect of sequential chemotherapy with Adr and Cp in 13762 RMTs. ● = control untreated tumors; ▲ = Cp (Day 14); △ = Adr (Day 14) + Cp (Day 16); ○ = Adr (Day 14) + Cp (Day 19); ■ = Cp (Days 14 and 23); □ = Adr (Day 14) + Cp (Days 15 and 24); ○ = Adr (Day 14) + Cp (Days 19 and 28). Each point = mean diameter ± 1 SE for at least 4 tumors expressed at a fraction of the Day-14 pretreatment size ($D_o$).

to 47 mg/m² in man (42) and is thus within a clinically applicable range. FLM analysis of unperturbed tumors showed the 13762 RMT to be a high-GF, rapidly proliferating tumor with high cell loss. In vitro methods for cell kinetic analysis provided similar results to those obtained by FLM analysis with fewer samples in a considerably shorter time. Previous studies in C3H/HeJ mammary tumors have also shown a good correlation between cell kinetic measurements made by in vitro and in vivo methods (32). Furthermore, the PDPI has been shown to estimate tumor GF in a wide range of experimental (34,36) and possibly human tumors (43).

After 5 mg/kg of Adr, $^3$H-TdR LIs were subnormal within 6 hours. Inasmuch as $T_S + TG_2$ was > 6 hours, the reduced $^3$H-TdR LI would indicate that at this dose level, Adr may affect the progression of cells through the S phase or $G_1$. The DLI value at 6 hours was, however, not significantly abnormal suggesting that at this time the movement of cells from S to $G_2$ phase was normal. $G_2$ progression delay has been reported in a number of systems after Adr (7,8,11,44). Although the methods used cannot detect a potential $G_2$ block, our results do suggest the possibility of $G_1$ block at this dose level.

While it is uncertain whether the PDPI accurately reflects changes in the GF in perturbed tumors, decreases in the PDPI during the first 3 days after Adr could be a reflection of a decreased proliferative fraction resulting from a dilution of the viable population by dead and/or doomed cells. Alternatively, since Adr binds to DNA and may remain for some time, the reduced PDPI may reflect decreased $^3$H-thymidine triphosphate incorporation during the PDP assay as a result of reduced template availability. If, however, the reduced PDPI was due simply to artifacts related to template inhibition without an at-

tendant decrease in cell proliferation, the $^3$H-TdR LI might be expected to remain at control levels. The Adr-induced kinetic perturbation suggested proliferative recovery 4–5 days after treatment. Cp treatment at Day 5, however, resulted in less tumor shrinkage than Cp treatment at 2 days.

Regrowth delay, perhaps a better measure of tumor response, after Cp was about 1–2 days longer when Cp was given on Day 5 after Adr. This relatively minor effect is not inconsistent with the relatively small changes in LI and PDPI between 2 and 5 days after Adr. Mean survival time was not different for animals treated with Cp either immediately after, 2 days, or 5 days after Adr. Although these survival time similarities cannot be ascribed to toxicity, they may be a manifestation of metastatic disease.

Metastatic incidence was not assessed in these studies, but previous reports have indicated established metastases within 18 days after 13762 tumor implantation (31).

Previous studies have indicated that Cp, given as two 80-mg/kg doses 9 days apart, is effective for inducing complete remission induction (80%) and some cures (30%). In the present studies initiation of this fractionated Cp schedule 5 days after Adr (a time corresponding to proliferative recovery) resulted in 100% remissions and cures. Fractionated Cp treatments beginning 1 day after Adr (a time corresponding to subnormal proliferation) was completely ineffective in producing long-term survivors.

In T1699 mammary tumors, proliferative recovery was noted 72 hours after 5 mg/kg of Adr. Sequential therapy studies with Adr and Cp in this model showed that a 72-sequencing interval was most effective (45). In L1210 leukemia, Adr and Cp have shown therapeutic synergism which may (16) or may not (17) be schedule-dependent. Other studies in a variety of animal models have shown the superiority of the two-drug combination over either drug given alone (10,18–23). Previous studies in our laboratory have shown that noncurative Cp schedules for 13762 RMTs could result in a high percentage of cures when Adr was added to the schedule at times predicted efficacious by changes in cell cycle kinetics indicative of proliferative recovery (19,20).

In advanced human breast cancer, patients without previous chemotherapy treated with oral Cp 3–6 days after Adr showed 87% objective and 22% complete response rates, with a 15-month median duration of complete response and 14-month overall median survival time (21).

The results from the present studies have shown that an Adr-induced cell kinetic perturbation can be detected using in vitro cell kinetic methods. Fractionated Cp treatments initiated during proliferative recovery after Adr were more effective than similar treatments initiated during intervals of subnormal proliferation. Although other factors such as drug distribution and pharmacodynamics could have been contributory, the results are consistent with and support the idea that efficacious treatment (chemotherapy and radiotherapy) sequencing intervals can be derived from the temporal changes in cell kinetic parameters after insult (10,19,20,27–29, 40,45). If applicable in other systems, this concept could represent a more rational approach to cancer therapy.

## REFERENCES


1. Arcamone F, Cassinelli G, Fantini G, et al. Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from *S. peucetius* var. *caesuis*. Biotechn Bioengin 11:1101–1110, 1969.
2. Arcamone F, Franceschi G, Penco S, et al. Adriamycin (14-hydroxydaunomycin), a novel antitumor antibiotic. Tetrahedron Lett 13:1007–1010, 1969.
3. Calendi E, DiMarco A, Reggiani M, et al. On physicochemical interactions between daunomycin and nucleic acids. Biochim Biophys Acta 103:25–49, 1965.
4. Di Marco A. Adriamycin (NSC-123127): mode and mechanism of action. Cancer Chemother Rep Part 3, vol 6(2):91–106, 1975.
5. Clarkson JM, and Humphrey RM. The effect of Adriamycin on cell cycle progression and DNA replication in Chinese hamster ovary cells. Cancer Res 37:200–205, 1977.
6. Bhuyan BK, and Fraser TJ. Cytotoxicity of antitumor agents in a synchronous mammalian cell system. Cancer Chemother Rep 58:149–155, 1974.
7. Barranco SC, Gerner EW, Burk KH, et al. Survival and cell kinetic effects of Adriamycin on mammalian cells. Cancer Res 33:11–16, 1973.
8. Hittleman WN, and Rao P. The nature of Adriamycin-induced cytotoxicity in Chinese hamster cells as revealed by premature chromosome condensation. Cancer Res 35:3027–3035, 1975.
9. Drewinko B, and Gottlieb JA. Survival kinetics of cultured human lymphoma cells exposed to Adriamycin. Cancer Res 33:1141–1145, 1973.
10. Schenken LL. Proliferative character and growth modes of neoplastic disease as determinants of chemotherapy efficacy. Cancer Chemother Rep 60:1761–1776, 1976.
11. Krishan A, and Frei E. Effect of Adriamycin on cell cycle traverse and kinetics of cultured human lymphocytes. Cancer Res 36:143–150, 1976.
12. Goldin A, and Johnson RK. Experimental tumor activity of Adriamycin (NSC-123127). Cancer Chemother Rep Part 3, vol 6(2):137–145, 1975.
13. Di Marco A, Lenaz L, Casazza AM, et al. Activity of Adriamycin (NSC-123127) and daunomycin (NSC-82151) against mouse mammary carcinoma. Cancer Chemother Rep 56:153–161, 1972.
14. Wasserman TH, Comis RL, Goldsmith M, et al. Tabular analysis of the clinical chemotherapy of solid tumors. Cancer Chemother Rep Part 3, vol 6(2):399–419, 1975.
15. Tormey DC. Adriamycin (NSC-123127) in breast cancer: an overview of studies. Cancer Chemother Rep Part 3, vol 6(2):319–327, 1975.
16. Tobias JS, Parker LM, Tattersall MHN, et al. Adriamycin/cyclophosphamide and Adriamycin/melphalan in advanced L1210 leukemia. Br J Cancer 32:199–207, 1975.
17. Avery TL, and Roberts D. Adriamycin and cyclophosphamide in combination chemotherapy of L1210 leukemia. Cancer Res 37:678–683, 1977.
18. Corbett TH, Griswold DP, Mayo JG, et al. Cyclophos-

phamide-Adriamycin combination chemotherapy of transplantable murine tumors. Cancer Res 35:1568–1573, 1975.
19. Braunschweiger PG, and Schiffer LM. Cell kinetics as a basis for chemotherapy scheduling in solid tumors. Proc Am Assoc Cancer Res and ASCO 17:57, 1976.
20. Braunschweiger PG, and Schiffer LM. Therapeutic implications of cell kinetic changes after cyclophosphamide treatment in "spontaneous" and "transplantable" mammary tumors. Cancer Treat Rep 62:727–736, 1978.
21. Jones SE, Durie BGM, and Salmon SE. Combination chemotherapy with Adriamycin and cyclophosphamide for advanced breast cancer. Cancer 36:90–97, 1975.
22. Lloyd RE, Jones SE, Salmon SE, et al. Combination chemotherapy with Adriamycin (NSC-123127) and cyclophosphamide (NSC-26271) for solid tumors: a phase II trial. Cancer Chemother Rep 60:77–83, 1976.
23. Soloway MS. Single and combination chemotherapy for primary murine bladder cancer. Cancer 36:333–340, 1975.
24. Schiffer LM. Prediction of drug time sequencing in vitro. Proc Am Assoc Cancer Res and ASCO 16:15, 1975.
25. Schiffer LM, Markoe AM, and Nelson JSR. Evaluation of the PDP index as a monitor of growth fraction during tumor therapy. *In* The Cell Cycle in Malignancy and Immunity. Oak Ridge, TN, Technical Information Center, USERDA, 1975, pp 459–472.
26. Schiffer LM, Braunschweiger PG, and Poulakos L. Rapid methods for utilizing cell kinetics for treatment in the C3H/He spontaneous mammary tumor: effects of vincristine. Cancer Chemother Rep 60:1913–1924, 1976.
27. Braunschweiger PG, and Schiffer LM. Cell kinetics after vincristine in spontaneous mammary tumors and implications for therapy. JNCI 60:1043–1048, 1978.
28. Braunschweiger PG, and Schiffer LM. Effect of methyl prednisolone on the cell kinetic response of C3H/HeJ mammary tumors to cyclophosphamide and Adriamycin. Cancer Res 39:3812–3815, 1979.
29. Braunschweiger PG, Schenken LL, and Schiffer LM. The cytokinetic basis for the design of efficacious radiotherapy protocols. Int J Oncol Biol Phys 5:37–47, 1979.
30. Segaloff A. Hormones and breast cancer. Recent Prog Horm Res 22:351–374, 1966.
31. Bodgen AE, Esber HJ, and Taylor DJ. Comparative study on the effects of surgery, chemotherapy and immunotherapy alone and in combination, on metastases of the 13762 mammary carcinoma. Cancer Res 34:1627–1631, 1974.
32. Braunschweiger PG, Poulakos L, and Schiffer LM. In vitro labeling and gold activation autoradiography for determination of labeling index and DNA synthesis times of solid tumors. Cancer Res 36:1748–1753, 1976.
33. Nelson JSR, and Schiffer LM. Autoradiographic detection of DNA polymerase containing nuclei in sarcoma 180 ascites cells. Cell Tissue Kinet 6:45–54, 1973.
34. Schiffer LM, Markoe AM, and Nelson JSR. Estimation of tumor growth fraction in murine tumors by the primer-available DNA-dependent DNA polymerase assay. Cancer Res 36:2415–2418, 1976.
35. Braunschweiger PG, and Schiffer LM. Nuclear DNA polymerase-α and replicative potential in mammalian cells. Eur J Cancer 13:775–779, 1977.
36. Braunschweiger PG, Poulakos L, and Schiffer LM. Cell kinetics in vivo and in vitro for C3H/He spontaneous mammary tumors. JNCI 59:1197–1204, 1977.
37. Wimber DE, and Quastler H. A $^{14}C$- and $^{3}H$-thymidine double labeling technique in the study of cell proliferation in *Tradescantia* root tips. Exp Cell Res 30:8–22, 1963.
38. Steel GG. Cell loss from experimental tumors. Cell Tissue Kinet 1:193–207, 1968.
39. Steel GG, and Hanes S. The technique of labeled mitoses: analysis by automatic curve-fitting. Cell Tissue Kinet 4:93–105, 1971.
40. Braunschweiger PG, Stragand JJ, and Schiffer LM. The effect of methylprednisolone on cell proliferation in C3H/HeJ spontaneous mammary tumors. Cancer Res 38:4510–4514, 1978.
41. Hopkins HA, Looney WB, Teja K, et al. Response kinetics of host and experimental solid tumors after Adriamycin. Br J Cancer 37:1006–1014, 1978.
42. Freireich EJ, Gehan EA, Rall DP, et al. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemother Rep 50:219–244, 1966.
43. Braunschweiger PG, and Schiffer LM. PDP indices in human tumors: evidence for proliferative correlations. Proc Am Assoc Cancer Res and ASCO 18:276, 1977.
44. Barlogie B, Drewinko B, Johnson DA, et al. The effect of Adriamycin on cell cycle traverse of a human lymphoid cell line. Cancer Res 36:1975–1979, 1976.
45. Braunschweiger PG, and Schiffer LM. T1699 tumor cure by cell kinetic-directed sequential chemotherapy. Proc Am Assoc Cancer Res and ASCO 20:79, 1979.

# *cis*-Dichlorodiammineplatinum(II) With and Without ICRF-159 in Non-Hodgkin's Lymphoma [1,2]

**Michael P. Corder,* Martin Wiesenfeld, Lawrence C. Maguire, J. Thompson Leimert, and Sherrie K. Panther [3]**

## SUMMARY

**Two studies of *cis*-dichlorodiammineplatinum(II) (CDDP) were conducted. The single-agent phase II study revealed that three of ten patients with non-Hodgkin's lymphoma treated with CDDP alone achieved objective partial remission of 7–15 weeks' duration. Eight patients were treated with CDDP plus ICRF-159. None of these eight patients achieved an objective response. CDDP is active as a single agent in non-Hodgkin's lymphoma. The combination of CDDP plus ICRF-159 does not appear to be a promising salvage program for non-Hodgkin's lymphoma.**

[Cancer Treat Rep 64:301–304, 1980]

*cis*-Dichlorodiammineplatinum(II) (CDDP) is a member of a new class of antineoplastic agents in man (1). In 1976, as part of a larger phase II trial, we began our study of this drug as a single agent. Substantial activity has been identified for this agent in Hodgkin's lymphoma (2,3). We have studied 18 patients with non-Hodgkin's lymphoma. Ten of these patients were treated with CDDP as a single agent and eight were treated with CDDP plus oral ICRF-159. ICRF-159 was selected for combination with CDDP because of (*a*) the demonstration of its activity against non-Hodgkin's lymphoma as a single agent in man (4), and (*b*) the observation of its additive effect in L1210 leukemia when combined with CDDP (5). The results of these two clinical trials are the subject of this report.

## PATIENTS AND METHODS

A broad phase II trial was initiated using CDDP on an outpatient schedule which used vigorous hydration and mannitol-induced diuresis as follows: 500 ml/hour of 5% dextrose in ½ N saline iv for 4 hours; 12.5 g of mannitol by iv push; 40 mg/m² of CDDP in 250 ml of normal saline over ½ hour; and 12.5 g of mannitol by iv push followed by 2½ hours of hydration at 500 ml/hour of 5% dextrose in ½ N saline. Initially, this program was scheduled to be given weekly with escalation of the dose to 55 mg/m² after four doses. Due to dose-limiting myelosuppression observed in the first 20 patients (6), the program was modified in November 1976 to 40 mg/m² iv for 3 weeks (2 weeks with no treatment), and if no hematologic toxicity was observed, 55 mg/m² was given for 3 weeks (2 weeks with no treatment). Of the ten patients who received the drug as a single agent three were receiving the former schedule and seven were receiving the latter schedule.

The schedule of CDDP plus ICRF-159 was as follows: CDDP was administered as above except that the dose was 55 mg/m² on Days 1 and 8 of a 35-day cycle; ICRF-159 was administered at a dose of 400 mg/m² orally every 8 hours for two doses on Days 2, 9, 16, 23, and 30 of the 35-day cycle.

There were no dose modifications for CDDP. If blood cell counts did not permit a scheduled dose, the dose was delayed until hematologic parameters permitted full doses (polymorphonuclear [PMN] count ≥ 2000/mm³ and platelet count ≥ 125,000/mm³). The ICRF-159 doses were modified according to the dose modification scheme shown in table 1.

Weekly observations included cbc, platelet count, creatinine, BUN, electrolytes (before and after infusion),

[1]Received Sept 24, 1979; revised Nov 21, 1979; accepted Nov 27, 1979.

[2]Supported in part by a grant from the Iowa Chapter of the Veterans of Foreign Wars Auxiliary, and by Public Health Service grant CA-18017 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare. Drs. Wiesenfeld and Maguire are Clinical Fellows of the American Cancer Society.

[3]Department of Internal Medicine, University of Iowa, Iowa City.

**Reprint requests to:* Michael P. Corder, MD, Section of Oncology, Department of Internal Medicine, University of Iowa, Iowa City, IA 52242.

TABLE 1.—Dose modifications for ICRF-159

| Wbc count (cells/mm³) | | Platelet count (cells/mm³) | Dose |
|---|---|---|---|
| If ≥ 4000 total | and | ≥125,000 | 100% |
| | *or* | | |
| If ≥ 2000 PMN | and | ≥125,000 | 100% |
| If 1500–1999 PMN | and/or* | 75,000–100,000 | 50% |
| If 1000–1499 PMN | and/or* | 50,000–75,000 | 25% |
| If < 1000 | or | <50,000 | Withhold dose |

*In case of discordance of PMN and platelet decision, the lower of the 2 dose levels is administered.

TABLE 2.—Patient characteristics and response to single-agent CDDP

| Patient No. | Histologic type* | Initial Karnofsky performance score | Response category† | Duration of response (wks) | Agents previously received‡ |
|---|---|---|---|---|---|
| 1 | DPDL + leukemic transformation | 50% | NR | — | Anthracycline antibiotics, alkylating agents, nitrosoureas, glucocorticoids, vinca alkaloids, ARA-C, M, MP |
| 2 | DH | 70% | PR | 15 | Anthracycline antibiotics, alkylating agents, nitrosoureas, glucocorticoids, vinca alkaloids, B, DTIC |
| 3 | NPDL + leukemic transformation | 70% | NR | — | Anthracycline antibiotics, alkylating agents, nitrosoureas, glucocorticoids, vinca alkaloids, B, total-body irradiation |
| 4 | WDLL → DH§ | 50% | NR | — | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, irradiation |
| 5 | NPDL → DH | 60% | NR | — | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, M |
| 6 | WDLL → DPDL | 40% | NR | — | Anthracycline antibiotics, alkylating agents, nitrosoureas, glucocorticoids, vinca alkaloids, B, PCZ, irradiation |
| 7 | DPDL (lymphoblastic) | 90% | NR | — | Anthracycline antibiotics, alkylating agents, vinca alkaloids, L-ASP, M, MP |
| 8 | UC | 60% | PR | 9 | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids |
| 9 | WDLL → DH | 60% | PR | 7 | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, B, intrathecal M, total-body irradiation |
| 10 | NM → DH | 50% | NR | — | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids |

*DPDL = diffuse poorly differentiated lymphocytic lymphoma; DH = diffuse histiocytic lymphoma; NPDL = nodular poorly differentiated lymphocytic lymphoma; WDLL = well-differentiated lymphocytic lymphoma; UC = unclassifiable non-Hodgkin's lymphoma; NM = nodular mixed lymphoma.

†PR = objective partial response; NR = less than an objective partial response.

‡ARA-C = cytosine arabinoside; B = bleomycin; M = methotrexate; L-ASP = L-asparaginase; MP = 6-mercaptopurine; PCZ = procarbazine.

§Histopathologic conversion.

TABLE 3.—Patient characteristics and response to CDDP + ICRF-159*

| Patient No. | Histologic type | Initial Karnofsky performance score | Response category | Agents previously received |
|---|---|---|---|---|
| 1 | DPDL + leukemic transformation | 50% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, irradiation. |
| 2 | DPDL | 80% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids |
| 3 | Nodular + diffuse mixed lymphomas | 70% | NR | Alkylating agents, glucocorticoids, vinca alkaloids, total-body irradiation |
| 4 | NPDL | 80% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, B |
| 5 | NPDL | 70% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, DTIC |
| 6 | NPDL | 40% | NR | Alkylating agents, glucocorticoids, vinca alkaloids, B |
| 7 | DPDL | 90% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids, B |
| 8 | DPDL | 90% | NR | Anthracycline antibiotics, alkylating agents, glucocorticoids, vinca alkaloids |

* *See* table 2 for definitions of abbreviations.

urinalysis, and tumor measurements. Pretreatment audiograms and bone marrow evaluations were obtained in all but one of these patients and were repeated when subjective complaints of eighth nerve damage or unexplained cytopenias occurred respectively.

The criteria of the Eastern Cooperative Oncology Group were used to assess response. A partial response could not be recorded without a $\geq$ 50% decrease in the sum of the products of the bidimensional measurable disease.

The total duration of response (table 1) was measured from the start of a $\geq$ 25% decrease in the measurable disease to the start of progression. Progression was defined as a $\geq$ 25% increase in the measurable disease.

All patients were treated according to the above protocol which had been approved by the University of Iowa Committee A (Human Investigation). Written informed consent was obtained from all patients.

## RESULTS

The response data and many of the patient characteristics for the two studies are shown in tables 2 and 3. In the single-agent study, six of nine evaluable patients had marrow involvement and nine of ten patients had stage IV disease.

Three of the ten patients achieved an objective response to CDDP when it was administered as a single agent. For these three patients, the time to start of response was 1, 1, and 4 weeks. Patients 1, 2, and 3 had received CDDP on a weekly schedule. Toxicity was primarily hematologic and gastrointestinal. Mean neutrophil and platelet count nadirs revealed a –54% and –33% change respectively from the baseline in evaluable patients (those patients [total, seven] not entering the study with neutropenia and/or thrombocytopenia). One patient had a neutrophil count nadir of $< 750/mm^3$ and two patients had platelet count nadirs of $< 50{,}000/mm^3$. Grade 1–2 nausea and vomiting was seen in all patients. One episode of apparent immediate hypersensitivity was observed and successfully re-treated with diphenhydramine prophylaxis as previously reported (7). Tinnitus was observed as a subjective complaint.

All eight patients who received CDDP plus ICRF-159 failed to achieve an objective response. Seven of these patients had extranodal disease at the time of therapy and six were demonstrated to have marrow involvement. The mean neutrophil and platelet count nadirs revealed a –75% and –81% change from the baseline in these patients. Neutropenia ($< 750$ cells/$mm^3$) was noted in four patients and thrombocytopenia ($< 50{,}000$ cells/$mm^3$) was noted in three. Seven of the eight patients experienced nausea and vomiting: one grade 1, five grade 2, and one grade 3. One of the 18 patients studied developed a transient elevation of creatinine to 2.0 mg/dl.

No serious electrolyte imbalance, drug-related sepsis, or significant bleeding developed in either study.

## DISCUSSION

Table 4 shows the reported patients with non-Hodgkin's lymphoma treated with CDDP. The overall response rate of 24% from the pooled series indicates that CDDP is an active agent in this diverse group of malig-

TABLE 4.—Response of non-Hodgkin's lymphoma to CDDP

| No. of Patients | No. of Objective responses | Authors (ref No.) |
|---|---|---|
| 17 | 1 | Rossof et al (3) |
| 5 | 3 | Talley et al (8) |
| 14 | 4 | Hill et al (9) |
| 10 | 3 | Corder et al (this paper) |
| 46 | 11 (24%) | |

nancies. Assuming a binomial distribution, the 95% confidence interval for the pooled response rate is 13%–39%. Although more hematologic toxicity was seen in these studies than is usually associated with CDDP, the fact that these patients are very heavily treated by any standards and also have a high propensity to bone marrow involvement probably accounts for this observation.

The trial in which we attempted to produce an increase in response rate of CDDP by the addition of ICRF-159, an agent with a reported 29% response rate in non-Hodgkin's lymphoma, was clearly a failure. The observation of no partial or complete responses in eight patients indicates that the maximum response rate, with a 95% confidence interval, would not be > 37% for the combination used in this setting. Our data also indicate that formidable hematologic toxicity is seen in these patients, with an indication that it *may* be significantly greater with the combination than with CDDP alone. The schedule of both CDDP and ICRF-159 was similar to that which has been associated with response rates of > 20% for our patients with non-Hodgkin's lymphoma. In order to justify intensive combination chemotherapy, the program should have the potential for substantially exceeding known single-agent response rates. This combination does not achieve this level of activity and further exploration of this combination as a "salvage" program is not warranted.

CDDP appears to be useful as a single agent in non-Hodgkin's lymphoma. The combination of CDDP with ICRF-159 produced significant toxicity without therapeutic benefit when administered by this dose and schedule.

## REFERENCES


1. ROZENCWEIG M, VON HOFF DD, SLAVIK M, ET AL. *cis*-Diamminedichloroplatinum (II). A new anticancer drug. Ann Intern Med 86:803–812, 1977.
2. CORDER MP, ELLIOTT TE, MAGUIRE LC, ET AL. Phase II study of *cis*-dichlorodiammineplatinum(II) in stage IVB Hodgkin's disease. Cancer Treat Rep 63:763–766, 1979.
3. ROSSOF AH, COLTMAN CA, JR, JONES SE, ET AL. Phase II evaluation of *cis*-dichlorodiammineplatinum(II) in lymphomas: a Southwest Oncology Group Study. Cancer Treat Rep 63:1605–1609, 1979.
4. FLANNERY EP, CORDER MP, SHEEHAN WW, ET AL. Phase II study of ICRF-159 in non-Hodgkin's lymphomas. Cancer Treat Rep 62:465–467, 1978.
5. WOODMAN RJ, SIRICA AE, GANG M, ET AL. The enhanced therapeutic effect of *cis*-platinum (II) diamminodichloride against L1210 leukemia when combined with cyclophosphamide or 1,2-bis(3,5-dioxopiperazine-1-yl)propane or several other antitumor agents. Chemotherapy 18:169–183, 1973.
6. CORDER MP, ELLIOTT TE, and BELL SJ. Dose limiting myelotoxicity in absence of significant nephrotoxicity with a weekly out-patient schedule of *cis*-platinum (II) diamminedichloride. J Clin Hematol Oncol 7:645–651, 1977.
7. WIESENFELD M, REINDERS E, CORDER M, ET AL. Successful re-treatment with *cis*-dichlorodiammineplatinum(II) after apparent allergic reactions. Cancer Treat Rep 63:219–221, 1979.
8. TALLEY RW, O'BRYAN RM, GUTTERMAN JU, ET AL. Clinical evaluation of toxic effects of *cis*-diamminedichloroplatinum (NSC-119875)—phase I clinical study. Cancer Chemother Rep 57:465–471, 1973.
9. HILL JM, LOEB E, PARDUE AS, ET AL. Platinum combination compounds in the treatment of acute leukemia and other malignant diseases with particular reference to malonato 1,2-diamino-cyclohexane platinum (II). J Clin Hematol Oncol 7:681–700, 1977.

# Complications of Peritoneovenous Shunt for Malignant Ascites [1,2]

J. Lokich,* R. Reinhold, M. Silverman, and J. Tullis [3]

## SUMMARY

**A peritoneovenous (PV) shunt was placed in eight patients with intractable malignant ascites. The shunt successfully controlled the ascites in six patients. The median survival time of the entire group was 2 months, with one patient alive at 22 months. Two of seven patients had secondary shunt failure: one from an unknown cause which could not be corrected by revision and another which was corrected by revision following removal of psammoma bodies in the valve. The complications of the shunt included transient edema (four patients), transient intravascular coagulation (four patients), and fever (two patients). Tumor embolization was suspected pathologically in two of eight patients although the ascitic fluid contained malignant cells in seven of eight patients. The PV shunt is a satisfactory palliative procedure for malignant ascites in the presence of adequate cardiac function and in the absence of urinary obstruction. The presence of bloody effusion or major intra-abdominal mass lesions contra-indicates a successful PV shunt. The acute adverse effects of the PV shunt (fever, fluid overload, and fulminant disseminated intravascular coagulation) may be prevented or minimized by preoperative fluid removal to obviate a major intravascular infusion of colloid and biologically active pyrogen and thromboplastin.**

[Cancer Treat Rep 64:305–309, 1980]

Neoplasia is a common cause of peritoneal effusions. Malignant ascites develops in > 50% of patients with ovarian cancer and in 10%–15% of patients with gastrointestinal cancer. The significant clinical manifestations of ascites include abdominal distention, pain, and inanition secondary to sequestration of protein-rich fluid. Palliation of malignant ascites by radiation therapy or intracavitary instillation of cytotoxic drugs or radioisotopes are effective in less than one half of the patients. Control of fluid accumulation is generally short-lived (1–3).

LeVeen et al (4,5) have reported the development of a peritoneovenous (PV) shunt with a one-way valve for the management of ascites secondary to cirrhosis. Anecdotal reports of the application of the LeVeen shunt to the management of ascites secondary to neoplastic disease have been cited by LeVeen et al (5), Arnot and White (6), Pollock (7), and, recently, by Straus et al (8). The complications of PV shunts have been infrequently addressed in terms of pathophysiology, and the specific criteria of eligibility for PV shunts in patients with malignant ascites are not established.

Our experience with the PV shunt in eight patients with intractable malignant ascites led to the development of guidelines for patient selection and determination of the specific role of the PV shunt in neoplastic disease.

## MATERIALS AND METHODS

Eight patients with malignant ascites secondary to ovarian cancer (three patients), breast cancer (three patients), mesothelioma (one patient), and an unknown primary cancer (one patient) were selected for shunt place-

[1]Received Sept 28, 1979; accepted Nov 26, 1979.

[2]The coagulation studies in this report were supported in part by grants from the Matilda R. Wilson Fund and the Rebecca Bernstein Fund.

[3]Departments of Medicine (Division of Medical Oncology), Surgery, and Pathology, New England Deaconess Hospital and Harvard Medical School, Boston, MA.

[4]We are grateful to Drs. J. Mc Caffrey and L. Cloud of the New England Deaconess Hospital staff for permission to include patients for study.

*Reprint requests to:* Jacob Lokich, MD, Division of Medical Oncology, New England Deaconess Hospital, Harvard Medical School, Boston, MA 02215.

ment. All patients had tumors which were resistant to standard forms of systemic tumor therapy and none had received abdominal radiation therapy or intracavitary radioisotopes. Four patients had received intracavitary chemotherapy without success. In six patients the ascites was associated with diffuse peritoneal implants of tumor (but without palpable mass lesions) and in two patients massive replacement of the liver by tumor was the dominant disease pattern. All patients had adequate cardiac and renal function prior to shunt placement without evidence of congestive heart failure or elevated serum creatinine.

Routine pathologic examination of the fluid included cytologic study in addition to cell-block analysis. Procoagulant activity was determined for both the supernatant and cell sediment of the ascitic fluid by employing standard assays for the prothrombin time, partial thromboplastin time, activated partial thromboplastin time, and thromboplastin generation time. Saline and plasma controls were determined concomitant with the evaluation of the ascitic fluid.

Peripheral blood was assayed to evaluate the systemic coagulation system employing the plasma prothrombin time, partial thromboplastin time, antithrombin III levels, and fibrin monomer and fibrin-split products. Determinations of systemic coagulation profiles were made on the day of surgery and daily for 3 days following surgery.

The majority of patients had lymphoscintography 1 day following surgery to determine shunt patency. Technetium 99m ($Tc^{99m}$) sulphate colloid (10 millicuries) was introduced into the abdominal cavity and gamma camera images of the upper thorax and abdomen were obtained at 1–4 hours.

The surgical technique for placement of the shunt has been described in detail by LeVeen et al (4,9) and by Reinhardt and Stanley (10). Under local or general anesthesia an incision is made in the abdominal wall and the peritoneal limb of the shunt is inserted into the abdominal cavity. The ascitic fluid is removed prior to connecting the shunt in order to obviate the potential complications of the infusion of large volumes of ascitic fluid into the vascular system acutely. The subcutaneous catheter is tunneled along the anterior or lateral chest wall and secured into the external or the internal jugular vein. The one-way valve is placed under the abdominal muscles extraperitoneally in a comfortable position.

The success of the PV shunt in managing ascites is based on the mechanical principle of the differential pressure between the abdominal cavity and the intrathoracic space. The one-way valve promotes flow from the abdominal cavity to the venous system as a consequence of respiratory motion, which simultaneously increases abdominal pressure by the downward movement of the diaphragm and decreases intrathoracic pressure by creating a negative pressure with a minimum differential of 5 cm of water.

## RESULTS

All patients had relief of ascites following the PV shunt. One patient died at 4 days, possibly secondary to a direct complication of the shunt, and is described in detail below. Two patients developed secondary shunt failure: one at 2 weeks with unsuccessful shunt revision and one at 3 months with successful shunt revision following removal of psammoma bodies from the valve. The median survival time of the seven patients who left the hospital was 7 months and two patients are alive, one at 22 months. The median survival time for all eight patients treated was only 2 months.

The clinical complications of the PV shunt are listed in table 1. Four patients developed abnormal peripheral coagulation profiles suggestive of intravascular coagulation, but in only one patient did this coagulopathy possibly contribute to the clinical picture.

## CASE REPORT

### Patient 2

Patient 2, a 70-year-old woman with papillary adenocarcinoma of the ovary, was treated with systemic chemotherapy for stage III disease from May to September 1977, when she developed progressive intractable malignant ascites requiring weekly paracentesis for relief of abdominal distention. She had a PV shunt placed and 4 hours after her operation developed severe pulmonary edema which responded partially to medical management. Prior to shunt placement the coagulation studies were normal. A coagulation profile 2 days after shunt demonstrated a prolonged prothrombin time and partial thromboplastin time, in addition to increased levels of fibrin-split products. Her platelet count and fibrinogen level, however, remained normal. She died on the fourth postoperative day of progressive respiratory insufficiency. At post-mortem examination, diffuse fibrin thrombi within the pulmonary arterioles were noted. In addition, the heart contained a mural thrombus in both the right and left ventricles and tumor cells were observed in the alveoli of the lung as well as in the thrombi of the left ventricle and in the pulmonary arterioles.

### Comment

The thrombi within the lung may have been a consequence of emboli from the right ventricular mural thrombus, but the possibility of ascitic fluid precipitating the clinical event cannot be excluded. The presence of tumor cells within the lungs and in the left ventricular mural thrombus may have been a consequence of intravascular infusion of tumor cells.

In six patients the pre-shunt peripheral coagulation screen was normal. In two patients circulating fibrin deg-

TABLE 1.—Clinical parameters and complications of PV shunt for malignant ascites in 8 patients

| Patient No. | Primary tumor | Cytology* | Systemic coagulation profile | | Complications | Survival |
|---|---|---|---|---|---|---|
| | | | Before shunt | After shunt | | |
| 1 | Ovary | + | Normal | Normal | None | 6 wks |
| 2 | Ovary | + | Normal | ↑Prothrombin time<br>↑Fibrin degradation products | Pulmonary embolism, tumor implants | 4 days |
| 3 | Ovary | + | ↑Fibrin degradation products | No change | Edema, secondary shunt failure | 7 mos |
| 4 | Breast | + | ↑Fibrin monomers | ↓Fibrinogen | Edema, pleural effusion, secondary shunt failure | 8 mos |
| 5 | Breast | – | Normal | Normal | Edema, fever | 28 mos |
| 6 | Breast | + | Normal | ↓Fibrinogen<br>↑Fibrin degradation products<br>↑Fibrin monomers | Fever, effusion | 1 mo† |
| 7 | Mesothelioma | +/– | Normal | ↓Fibrinogen<br>↑Prothrombin time<br>↑Fibrin degradation products | Tumor implants | 7 mos |
| 8 | Tumor of unknown origin | + | Normal | No change | Edema, pleural effusion | 2 mos |

*+ = positive; – = negative.
†Still alive.

radation products or fibrin monomers were observed. In all patients tested, the ascitic fluid analysis demonstrated the presence of thromboplastic material in either the supernatant, the sediment, or both. Details on the methodology and clinical implications of identifying thromboplastic activity will be reported separately. Comparative quantification of the procoagulant activity was not feasible in this small group of patients, but the amount of thromboplastic activity of the ascitic fluid did not always predict for the development of subsequent peripheral coagulation abnormalities. In all patients the observed coagulation abnormality following shunt placement generally suggested intravascular coagulation, but the condition was transient with restitution to normal coagulation parameters by Day 3.

A second common complication of PV shunt was edema of the peripheral extremities (four patients) with or without pleural effusion related to the intravascular expansion with the ascitic infusate. In three of the four patients the edema receded gradually over a 2–3-week period. In one of the four patients, edema was a persistent complication. This patient required twice monthly paracentesis to retain a functional life style, but the function of the shunt was maintained as shown by $Tc^{99m}$ sulphate lymphoscintography (fig 1). In addition, the patient demonstrated an improved nitrogen balance as shown by increased muscle strength attributed to salvage and utilization of the ascitic fluid protein.

Fever was observed in two patients. One patient developed spiking fevers daily following shunt placement, but blood and other cultures were negative and sepsis was not documented. Seven days after shunt placement the fever subsided. The mechanism of the fever is unknown, but the infusion of a pyrogen in the ascitic fluid is one possibility.

## DISCUSSION

LeVeen et al (4) have reported 18 patients with malignant ascites in whom the shunt was placed in the terminal states, but specific details of the effectiveness of the shunt were not indicated. Arnot and White (6) reported two of three patients in whom PV shunts relieved malignant ascites, and Pollock (7), using a Holter valve in a PV shunt in nine patients with malignant ascites, produced a single long-term survival. Straus et al (8) reported 37 patients in whom a PV shunt was performed for malignant ascites (predominantly gynecologic cancer); in 27 (73%) of these patients, the ascites was successfully controlled, but the median survival time of the group was only 8 weeks. More recent reports by Wilbanks et al (11) and by Qazi et al (12) have summarized more limited experiences. In the study by Wilbanks et al, 17 of 20 patients with gynecologic malignancies (predominantly

FIGURE 1.—$Tc^{99m}$ scan of thorax obtained 1 hr after intra-abdominal instillation demonstrating the contour of the shunt catheter.

ovary) had a successful shunt without complications noted. Qazi et al similarly studied 20 patients (14 with ovarian cancer), and in 12 (65%) of these patients palliation was effected without significant complications.

In our study the PV shunt provided meaningful palliation in six of eight patients, all of whom had relief of ascites. The complications of PV shunts have not been appreciated in the past but are important in the risk-benefit determination and patient selection. Candidates for PV shunts must undergo an evaluation of cardiac, pulmonary, and renal function and reserve in that additional volume load on these systems may lead to cardiopulmonary and renal failure. Other contra-indications to PV shunt include the presence of large intra-abdominal mass lesions or bloody peritoneal effusion. In one series only one of eight PV shunts functioned with bloody ascites (8).

Complications of the PV shunt will still occur in selected patients, and the types of complications experienced in this series are typical. Disseminated intravascular coagulation has been previously reported by LeVeen (9) and has been fatal in another report by Matseshe et al (13), both in patients with nonmalignant ascites. The development of coagulopathy in patients re-infused with ascitic fluid is common (14,15), but the specific mechanism has not been extensively studied. Harmon et al (16) have, in fact, in a retrospective study of 19 shunts, suggested that disseminated intravascular coagulation develops in all patients in whom the shunt is successful and therefore is a means of evaluating shunt function.

Recently, analysis of ascitic fluid has demonstrated thromboplastic material and mucin-secreting tumors which are especially rich in thromboplastic material (17). In patient 2, thromboplastic activity was demonstrated in the ascitic fluid and may have been a primary cause of fibrin deposition within the pulmonary arterioles. Thromboplastic material in the ascitic fluid may be in the sediment (cellular) fraction or the supernatant or both. The severity of the coagulopathy which develops is variable and was of potential significance in only one patient in this report. The clinical significance of the coagulopathy may be diminished in part by the practice of removing the majority of fluid prior to activation of the shunt, thus precluding an acute overwhelming coagulopathy.

Tumor embolization is the second important potential complication of PV shunts. In two patients in this series tumor cells were observed at autopsy within the lungs and the pattern of implantation was pathologically suggestive of embolization. In one patient reported by LeVeen and co-workers, tumor implantation on the superior vena cava was observed. Straus et al failed to demonstrate tumor implants in nine patients who were autopsied. Although the ascitic fluid contained malignant cells in seven of eight patients, only two patients in our study had suggestive clinical or pathologic evidence of embolization. Since the PV shunt is a palliative procedure in preterminal patients with widely disseminated tumor, the post-mortem examination cannot readily distinguish disbursement and implantation of cells by the shunt from natural hematogenous metastases.

In summary, the PV shunt developed by LeVeen is an effective palliative measure in the management of malignant ascites. Candidates for the shunt should demonstrate adequate cardiac and renal function in order to tolerate the large fluid loads imposed by the additional burden of protein-rich fluid. An evaluation of the procoagulant properties of the ascitic fluid did not prove to be of value in identifying patients who develop clinically significant coagulopathy. Although the ascitic fluid was uniformly positive for thromboplastic activity and many patients also demonstrated alteration of blood coagulation profiles following shunt, they did not necessarily develop sustained intravascular coagulation. Removal of the ascitic fluid prior to the surgical procedure may be an important step in preventing significant acute adverse effects of the shunt. The use of lymphoscintography postoperatively is a helpful method of establishing patency of the shunt (18).

## REFERENCES


1. Card RY, Cole DR, and Henschke UK. Summary of ten years of the use of radioactive colloids in intracavitary therapy. J Nucl Med 1:195–198, 1960.
2. Strober JS, Klatz E, and Kaperwau A. Malignant pleural disease: a radiotherapeutic approach to the problem. JAMA 226:296–299, 1973.
3. Groesbeck HP, and Cudmore JTP. Intracavitary thio-TEPA for malignant effusions. Am Surg 28:90–95, 1962.
4. LeVeen HH, Christoudias G, IP M, et al. Peritoneo-venous shunting for ascites. Ann Surg 180:580–599, 1974.
5. LeVeen HH, Wapnick S, Grosberg S, et al. Further experience with peritoneo-venous shunt for ascites. Ann Surg 184:574–581, 1976.
6. Arnot RS, and White H. Letter: LeVeen shunts. Lancet 1(8062):505, 1978.
7. Pollock AV. The treatment of resistant malignant ascites by insertion of a peritoneo-atrial Holter valve. Br J Surg 62:104–107, 1975.
8. Straus AK, Roseman DL, and Shapiro TM. Peritoneovenous shunting in the management of malignant ascites. Arch Surg 114:489–491, 1979.
9. LeVeen HH. Peritoneo-jugular shunt for ascites. *In* Resident and Staff Physician, 1978, pp 96–106.
10. Reinhardt GF, and Stanley MM. Peritoneovenous shunting for ascites. Surg Gynecol Obstet 145:419–424, 1977.
11. Wilbanks GD, Straus AK, Roseman DL, et al. Peritoneovenous shunting in the management of refractory ascites from gynecologic malignancy. Proc Am Assoc Cancer Res and ASCO 20:364, 1979.
12. Qazi R, Savlov ED, and Sischy B. Peritoneo-venous shunt for palliation of malignant ascites. Proc Am Assoc Cancer Res and ASCO 20:403, 1979.
13. Matseshe JW, Beart RW, Jr, Bartholomew LG, et al. Fatal disseminated intravascular coagulation after peritoneovenous shunt for intractable ascites. Mayo Clin Proc 53:526–528, 1978.
14. Ansley JD, Bethel RA, Bowen A II, et al. Effect of peritoneovenous shunting with the LeVeen valve on ascites, renal function, and coagulation in six patients with intractable ascites. Surgery 83:181–187, 1978.
15. Villeneuve JP, Thuot C, Marleau D, et al. Treatment of resistant ascites by continuous ultrafiltration—reinfusion of ascitic fluid. Can Med Assoc J 117:1296–1298, 1977.
16. Harmon DC, Demirjian Z, Ellman L, et al. Disseminated intravascular coagulation with the peritoneovenous shunt. Ann Intern Med 90:774–776, 1979.
17. Lerner RG, Nelson JC, Corines P, et al. Intravascular coagulation complicating peritoneal-atrial shunts. Proc Am Soc Hematol 50:274, 1977.
18. Kirchmer N, and Hart U. Radionuclide assessment of LeVeen shunt patency. Ann Surg 185:145–146, 1977.

# Lack of Clinical Therapeutic Synergism Between Cyclophosphamide, Adriamycin, and *cis*-Dichlorodiammineplatinum(II) in Metastatic Human Colorectal Carcinoma [1,2]

Brief Reports and Preliminary Communications

**Michael J. O'Connell,* Charles G. Moertel, Joseph Rubin, Allan J. Schutt, and Edward T. Creagan [3]**

## SUMMARY

**Thirty patients with metastatic adenocarcinoma originating in the colon or rectum were treated with a combination of cyclophosphamide, Adriamycin, and *cis*-dichlorodiammineplatinum(II). Only one partial tumor regression of 12 weeks' duration was observed. Severe nausea and vomiting secondary to chemotherapy were frequently seen. This three-drug regimen, as employed in our study, was not associated with useful clinical activity against human metastatic colorectal cancer.**

[Cancer Treat Rep 64:311–313, 1980]

In our experience, cyclophosphamide, Adriamycin, and *cis*-dichlorodiammineplatinum(II) (cisplatin) (CAP), when used as single agents, have shown only minimal activity against human colorectal carcinoma, with transient objective responses being observed among five of 25 patients treated with cyclophosphamide (1), four of 57 treated with Adriamycin (2), and two of 51 treated with cisplatin (3). However, these agents have been reported to show therapeutic synergism when used in combination in some animal model systems (4,5). Furthermore, Eagan and associates (6) have reported significant clinical activity of these three drugs used in combination for the treatment of bronchogenic non-small cell lung cancer including adenocarcinomas, tumors which have historically been chemotherapy-resistant. We elected, therefore, to study the CAP regimen in patients with metastatic colorectal adenocarcinoma using the same doses of cyclophosphamide and Adriamycin which were successfully employed in the treatment of advanced adenocarcinoma of the lung. Cisplatin was administered at a higher dose level by 3-hour infusion with mannitol diuresis being added in the current study in an attempt to minimize vomiting and renal toxic effects.

## PATIENTS AND METHODS

### Patient Characteristics

All patients had histologic confirmation of unresectable metastatic adenocarcinoma originating from the colon or rectum. Each patient had at least one area of measurable metastatic disease to serve as an objective indicator of response to therapy.

All patients were ambulatory and were maintaining adequate oral nutrition. Informed consent was obtained. Patients were excluded if there was significant renal impairment (serum creatinine $> 2.0$ mg/100 ml), depression of peripheral blood cell counts (wbcs $< 4000/mm^3$ and/or platelets $< 130{,}000/mm^3$), frequent vomiting, active infection, gastrointestinal bleeding, or underlying organic heart disease.

Pretreatment patient characteristics are summarized in table 1.

### Treatment

All patients had a pertinent medical history and physical examination including measurement of indicator lesion(s), serum chemistry profile (including creatinine, alkaline phosphatase, SGOT, and bilirubin), wbc and platelet counts, hemoglobin level, chest roentgenogram, and urinalysis before initial therapy and before each subsequent treatment and an electrocardiogram before treatment.

Cyclophosphamide was administered at a dose of 400

[1]Received June 22, 1979; revised Nov 6, 1979; accepted Nov 15, 1979.

[2]Supported in part by contract N01-CM-02066 from the Division of Cancer Treatment, National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[3]Division of Medical Oncology, Mayo Clinic, Rochester, MN.

**Reprint requests to:* Dr. Michael J. O'Connell, Department of Oncology, Mayo Clinic, Rochester, MN 55901.

TABLE 1.—Pretreatment patient characteristics

| Characteristic | No. of patients |
|---|---|
| Prior therapy | |
| None | 15 |
| Chemotherapy only | 11 |
| Radiotherapy + chemotherapy | 4 |
| Performance status* | |
| 0–1 | 18 |
| 2–3 | 12 |
| Primary indicator lesion | |
| Liver | 19 |
| Lung | 7 |
| Abdominal mass | 2 |
| Peripheral lymph node | 2 |

*Eastern Cooperative Oncology Group scale: 0 = fully active to 4 = completely disabled.

mg/m² by rapid iv push. One hour later, Adriamycin was given at a dose of 40 mg/m² by iv push followed by a 3-hour iv infusion consisting of cisplatin at a dose of 60 mg/m² plus 25 g of mannitol diluted in 1500 ml of 5% dextrose in one half normal saline. The initial dose of cisplatin was reduced to 45 mg/m² if the pretreatment serum creatinine was 1.3–1.4 mg/100 ml (one patient) and was reduced to 30 mg/m² if the pretreatment serum creatinine was 1.5–1.9 mg/100 ml (one patient). Treatment was given on Day 1 only and was repeated at 28-day intervals for patients who experienced an objective response or remained stable. Chemotherapy dose was decreased as necessary in the event of severe vomiting, myelosuppression, or renal toxic effects.

### Response Criteria

A > 50% reduction in the sum of the products of bidimensionally measurable lesions or a > 30% reduction in the sum of linear measurements below the costal margin, if malignant hepatomegaly was the indicator lesion, was required to achieve an objective response. No new areas of malignant disease could occur. Furthermore, the observation must have been made at least 8 weeks following the start of therapy.

## RESULTS

### Therapeutic Results

One of the 15 patients previously treated with chemotherapy experienced a regression of hepatic metastases which lasted 12 weeks from the start of therapy. None of the previously untreated patients achieved an objective tumor response. Overall, of the 30 patients treated, 16 experienced tumor progression by 4 weeks after the first treatment and 25 had tumor progression by 8 weeks after the first treatment. The maximum number of cycles administered to any patient was four. The median survival time measured from the first day of therapy was 129 days for previously treated patients and 114 days for previously untreated patients.

### Toxicity

Chemotherapy toxicity is summarized in table 2. Occasionally, severe myelosuppression was observed. In addition, the majority of patients experienced clinically significant nausea, vomiting, and anorexia. Vomiting was considered severe in 15 of the patients, in spite of liberal use of antiemetics such as prochlorperazine. There was no evident renal toxicity.

## DISCUSSION

The lack of clinical value of the CAP regimen for the management of advanced colorectal carcinoma is evident on the basis of the very low objective response rate and considerable toxicity observed. The use of a 3-hour cisplatin infusion with mannitol diuresis did not prevent the significant nausea and vomiting which were seen in most patients. Unfortunately, we were not able to demonstrate an appreciable objective response rate as one would hope for if these agents were acting in a synergistic fashion.

Furthermore, we could not document significant antitumor effect against colorectal adenocarcinoma from the same combination of drugs which yielded objective tumor responses in six of 14 previously untreated patients with bronchogenic adenocarcinoma at our institution (6). More recent experience with the CAP regimen in the treatment of advanced bronchogenic adenocarcinoma, in which the

TABLE 2.—Toxic effects in 30 patients

| Toxic effect | No. of patients |
|---|---|
| Nausea | 30 |
| Vomiting | 29 |
| Anorexia | 16 |
| Stomatitis | 3 |
| Alopecia | 23 |
| Diarrhea | 7 |
| Wbc count nadir (cells/mm³) | |
| < 4000 | 22 |
| < 2000 | 8 |
| < 1000 | 3 |
| Platelet count nadir (cells/mm³) | |
| < 130,000 | 8 |
| < 50,000 | 1 |

dose of each chemotherapy drug was the same as that employed in the present study and in which cisplatin was given as a 6-hour infusion, indicated an objective response rate of 45% (ten of 22 previously untreated patients) (7). These observations underscore the variable chemotherapy sensitivity of adenocarcinomas which arise from different anatomic sites.

## REFERENCES


1. SCHUTT AJ, HAHN RG, REITEMEIER RJ, ET AL. A phase 2 study of intermittent high-dose cyclophosphamide therapy of advanced gastrointestinal cancer. Cancer Res 33:2218–2220, 1973.
2. FRYTAK S, MOERTEL CG, SCHUTT AJ, ET AL. Adriamycin (NSC-123127) therapy for advanced gastrointestinal cancer. Cancer Chemother Rep 59:405–409, 1975.
3. KOVACH JS, MOERTEL CG, SCHUTT AJ, ET AL. Phase II study of *cis*-diamminedichloroplatinum (NSC-119875) in advanced carcinoma of the large bowel. Cancer Chemother Rep 57:357–359, 1973.
4. CORBETT TH, GRISWOLD DP, MAYO JG, ET AL. Cyclophosphamide-Adriamycin combination chemotherapy of transplantable murine tumors. Cancer Res 35:1568–1573, 1975.
5. WOODMAN RJ, SIRICA AE, GANG M, ET AL. The enhanced therapeutic effect of *cis*-platinum (II) diamminodichloride against L1210 leukemia when combined with cyclophosphamide or 1,2-bis(3,5-dioxopiperazine-1-yl)propane or several other antitumor agents. Chemotherapy 18:169–183, 1973.
6. EAGAN RT, INGLE JN, FRYTAK S, ET AL. Platinum-based polychemotherapy versus dianhydrogalactitol in advanced non-small cell lung cancer. Cancer Treat Rep 61:1339–1345, 1977.
7. EAGAN RT, FRYTAK S, CREAGAN ET, ET AL. Phase II study of cyclophosphamide, Adriamycin, and *cis*-dicholorodiammineplatinum(II) by infusion in patients with adenocarcinoma and large cell carcinoma of the lung. Cancer Treat Rep 63:1589–1591, 1979.

# Evaluation of Vitamin E and Selenium Protection Against Chronic Adriamycin Toxicity in Rabbits [1,2,3]

Brief Reports and Preliminary Communications

John F. Van Vleet[4,*] and Victor J. Ferrans[5,6]


## SUMMARY

**Chronic Adriamycin (ADR) toxicity was produced in 50 weanling rabbits by weekly injections of 2.4 mg of ADR/kg of body weight for up to 17 weeks. An evaluation was made of the protective value of weekly injected supplements, given 24 hours before each ADR injection, which included: (*a*) low-dose vitamin E (α-tocopherol acetate, 17 mg/kg) (*b*) high-dose vitamin E (α-tocopherol acetate, 170 mg/kg); (*c*) selenium (as sodium selenite, 0.06 mg/kg); and (*d*) combined low-dose vitamin E and selenium. The incidence of cardiomyopathy was high in both supplemented and unsupplemented rabbits. Mean cardiomyopathy scores were low in the unsupplemented rabbits and moderately high in the supplemented rabbits (that survived longer), but were lower in rabbits given high-dose vitamin E than in the other supplemented groups. Supplements of vitamin E, selenium, or both prolonged survival of animals treated continuously with ADR, but only supplements of large amounts of vitamin E provided any evidence of even partial protection against chronic cardiotoxicity.**


[Cancer Treat Rep 64:315–317, 1980]

Chronic cardiotoxicity presents a major limitation to the usefulness of Adriamycin (ADR). Studies of cardioprotection with chemical antidotes that can be administered concurrently with ADR have given variable results. Large amounts of vitamin E protected against acute cardiotoxicity in mice (1). Partial protection against chronic cardiotoxicity was provided by coenzyme Q10 in rabbits (2), rats (3), and man (4), by large doses of vitamin E in rats (5), and by small doses of selenium and vitamin E in rabbits (6). Digitoxin, strophanthin-K, ouabain, and coenzyme Q7 were not cardioprotective in animals (7). The present study was designed to determine the value of supplements of vitamin E and selenium, given alone and in combination, against the development of chronic ADR cardiotoxicity in rabbits.

## MATERIALS AND METHODS

Sixty weanling male New Zealand white rabbits (body weight, 2 kg) were allotted to six equal groups (A–F). Rabbits in groups A–E were given weekly iv injections of 2.4 mg of ADR/kg of body weight. Rabbits in groups B–E were also given an im injection of either vitamin E (groups B and C), selenium (group D), or both (group E) suspended in the vehicle of the commercial Se-E preparation (L-Se, Burns-Biotec Laboratories) given to group E. Group B (low-dose vitamin E) received 17 mg/kg of α-tocopherol acetate; group C (high-dose vitamin E) received 170 mg/kg of α-tocopherol acetate; and group D (selenium) received 0.06 mg/kg of selenium as sodium selenite and 17 mg/kg of α-tocopherol acetate. Group F received weekly iv injections of physiologic saline solution in amounts comparable to those given to the ADR-treated rabbits. Blood samples were collected weekly for determinations of packed cell volume, hemoglobin concentration, and erythrocyte count. At 17 weeks all of the ADR-treated rabbits had died, and the control animals were killed and necropsied. The severity of hydrothorax, cardiomyopathy, and nephropathy was graded as shown in table 1. Statistical analysis by single-factor analysis of

[1]Received July 2, 1979; revised Oct 19, 1979; accepted Nov 26, 1979.

[2]Submitted as Journal Paper No. 7412, Purdue University Agricultural Experiment Station, West Lafayette, IN.

[3]Supported in part by a grant-in-aid from the American Heart Association, Indiana Affiliate, Inc.

[4]Department of Veterinary Microbiology, Pathology, and Public Health, School of Veterinary Medicine, Purdue University, West Lafayette, IN.

[5]Pathology Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

[6]We thank Dr. David M. Young, Laboratory of Toxicology, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute (Bethesda, MD), for supplying Adriamycin for this study and Burns-Biotec Laboratories, Division of Chromalloy Pharmaceutical, Inc (Oakland, CA), for providing the selenium and vitamin E preparations for this study.

*Reprint requests to:* John F. Van Vleet, DVM, PhD, Department of Veterinary Microbiology, Pathology, and Public Health, School of Veterinary Medicine, Veterinary Pathology Bldg, Purdue University, West Lafayette, IN 47907.

TABLE 1.—Effect of vitamin E, selenium, or both on survival time and severity and incidence of cardiomyopathy, hydrothorax, and nephropathy in rabbits with continuous ADR administration

| Group | Treatment | No. of rabbits | Mean survival time in days ± SD (minimum-maximum) | Mean total ADR dose ($mg/m^2$) | Mean cardiomyopathy severity scores* ± SD (minimum-maximum) and incidence of lesions (No. of rabbits affected/total rabbits) | Mean hydrothorax severity score† ± SD and incidence of lesions (No. of rabbits affected/total rabbits) | Mean nephropathy severity score‡ ± SD and incidence of lesions (No. of rabbits affected/total rabbits) |
|---|---|---|---|---|---|---|---|
| A | ADR | 10 | $62.7^{a}$§ ± 16.5 (39–98) | $245.5^{a}$ | $0.42^{a}$ ± 0.34 (0.00–0.93) (9/10) | $0.80^{a}$ ± 0.92 (5/10) | $2.20^{a}$ ± 0.79 (10/10) |
| B | ADR and low-dose vitamin E | 10 | $86.0^{b}$ ± 24.7 (50–119) | $335.3^{b}$ | $1.72^{b}$ ± 0.84 (0.13–2.87) (10/10) | $1.60^{a}$ ± 1.26 (7/10) | $2.30^{a}$ ± 0.82 (10/10) |
| C | ADR and high-dose vitamin E | 10 | $90.2^{b}$ ± 25.4 (49–122) | $348.5^{b}$ | $1.01^{ab}$ ± 0.80 (0.07–2.87) (10/10) | $1.40^{a}$ ± 1.17 (7/10) | $2.40^{a}$ ± 0.97 (10/10) |
| D | ADR and selenium | 9 | $90.7^{b}$ ± 18.8 (59–124) | $354.9^{b}$ | $1.85^{b}$ ± 1.39 (0.00–3.64) (8/9) | $1.56^{a}$ ± 1.51 (5/9) | $2.67^{a}$ ± 0.72 (9/9) |
| E | ADR, selenium, and low-dose vitamin E | 10 | $76.3^{ab}$ ± 21.6 (48–118) | $295.7^{ab}$ | $1.58^{b}$ ± 0.97 (0.00–2.93) (9/10) | $1.60^{a}$ ± 1.17 (8/10) | $2.00^{a}$ ± 0.94 (9/10) |
| F | Controls | 10 | — | — | — | — | — |

*Cardiomyopathy severity scores were graded by histologic study: 0 = none, 1 = focal, 2 = mild-diffuse, 3 = moderate-diffuse, 4 = severe-diffuse.
†Hydrothorax severity scores were graded by macroscopic examination: 0 = none, 1 = mild, 2 = moderate, 3 = severe.
‡Nephropathy severity scores were graded by histologic study: 0 = none, 1 = mild, 2 = moderate, 3 = severe.
§Means with different superscripts, (a,b) in same vertical column are significantly different ($P < 0.05$).

variance and Newman-Keuls test was applied to survival time, severity scores of hydrothorax, cardiac and renal lesions, and laboratory data to determine significant differences between treatment groups.

## RESULTS

Mean survival time was prolonged in ADR-treated rabbits receiving supplements of vitamin E, selenium, or both (groups B–E), but great variability was present within groups. Mean cardiomyopathy scores were low in unsupplemented ADR-treated rabbits (group A), high in ADR-treated rabbits given low-dose vitamin E (group B), selenium (group D), or both (group E), and intermediate in ADR-treated rabbits given high-dose vitamin E (group C); however, great variability was present within groups. The severity scores tended to be moderately dose-related, but frequent exceptions were seen. The incidence of cardiomyopathy was uniformly high in all groups. Mean hydrothorax severity scores and incidence of hydrothorax were greatest in supplemented, ADR-treated groups (groups B–E).

Mean nephropathy severity scores and incidence of nephropathy were high in all ADR-treated rabbits (groups A–E); however, these scores tended to be dose-related. Rabbits in groups B–E had delayed onset of severe renal disease. This tended to correlate with the prolonged survival also observed in these groups. Mean hemoglobin concentration, erythrocyte count, and packed cell volume showed a progressive decrease in all ADR-treated rabbits (groups A–E).

## DISCUSSION

The present study demonstrates prolongation of survival time in ADR-treated rabbits given supplements of vitamin E, selenium, or both. These results support the findings of our previous 10-week study showing decreased cumulative mortality in ADR-treated rabbits that received selenium together with vitamin E (6). In contrast to our previous study, the incidence and severity of cardiomyopathy in the present study were no greater in the unsupplemented than in the supplemented rabbits. In fact, although the unsupplemented rabbits died earlier, the mean cardiomyopathy scores in these rabbits were substantially lower than those in the supplemented rabbits, even when specifically compared to only those scores in the few supplemented rabbits that died after receiving similar small total amounts of ADR.

The group given the high-dose vitamin E supplement had a lower mean cardiomyopathy score than did the other three supplemented groups. This finding is in agreement with previous reports that high doses of vitamin E protect against acute (1) and chronic (5,6) ADR-induced cardiotoxicity. It is also important to recognize that although there was a trend toward higher cardiomyopathy scores as rabbits were given larger cumulative doses of ADR, there were many exceptions to this pattern. Some rabbits developed severe cardiomyopathy after receiving only small cumulative doses of ADR and others had little or no cardiac damage after receiving large doses.

In summary, supplements of selenium or vitamin E to ADR-treated rabbits offer only limited protection against the eventual development of chronic cardiomyopathy, a lesion that we consider to be mediated principally via nuclear injury and decreased synthesis of contractile proteins (8). Prolonged survival of rabbits supplemented with selenium or vitamin E may be the result of protection against early ADR-induced lipoperoxidative damage.

## REFERENCES


1. Myers CE, McGuire W, and Young R. Adriamycin: amelioration of toxicity by α-tocopherol. Cancer Treat Rep 60:961–962, 1976.
2. Bertazzoli C, Sala L, Solcia E, et al. Experimental Adriamycin cardiotoxicity prevented by ubiquinone in vivo in rabbits. Int Res Commun Syst Med Sci 3:468, 1975.
3. Zbinden G, Bachman E, and Bollinger H. Study of coenzyme Q in toxicity of Adriamycin. *In* Biomedical and Clinical Aspects of Coenzyme Q (Folkers K, and Yamamura Y, eds). Amsterdam, The Netherlands, Elsevier North-Holland Biomedical Press, 1977, pp 219–228.
4. Cortes EP, Gupta M, Chou C, et al. Adriamycin cardiotoxicity: early detection by systolic time interval and possible prevention by coenzyme Q10. Cancer Treat Rep 62:887–891, 1978.
5. Sonneveld P. Effect of α-tocopherol on the cardiotoxicity of Adriamycin in the rat. Cancer Treat Rep 62:1033–1036, 1978.
6. Van Vleet JF, Greenwood L, Ferrans VJ, et al. Effect of selenium-vitamin E on Adriamycin-induced cardiomyopathy in rabbits. Am J Vet Res 39:997–1010, 1978.
7. Philips FS, Gilladoga A, Marquardt H, et al. Some observations on the toxicity of Adriamycin (NSC-123127). Cancer Chemother Rep Part 3, vol 6(2):177–181, 1975.
8. Ferrans VJ. Overview of cardiac pathology in relation to anthracycline cardiotoxicity. Cancer Treat Rep 62:955–961, 1978.

# Complete Remission of Mediastinal Germ-Cell Tumors With *cis*-Dichlorodiammineplatinum(II) Combination Chemotherapy [1]

Brief Reports and Preliminary Communications

Q. G. C. M. van Hoesel[2,*] and H. M. Pinedo[2,3,4]

## SUMMARY

**Two patients with mediastinal germ-cell tumors achieved a complete remission with the combination of vinblastine, bleomycin, and *cis*-dichlorodiammineplatinum(II).**

[Cancer Treat Rep 64:319–321, 1980]

The occurrence of germ-cell tumors in the mediastinum is well-known and has given rise to speculation about their origin. The question is whether these neoplasms arise as primary tumors in localizations other than the testes.

Because fibrous scars (1,2) and microscopic tumor foci (3) are found in the testes of patients with extensive extratesticular neoplasms at autopsy, the clinician must decide whether a mediastinal tumor with histologic features of testicular carcinoma is a primary or a metastatic lesion. Utz and Buscemi (4) consider this neoplasm to be a primary tumor if the testes are normal on palpation. The absence of retroperitoneal disease may provide support for this conclusion.

The prognosis of mediastinal tumors is not determined solely by the response to chemotherapy. Patients may die shortly after presentation due to invasion or compression of vital organs. Treatment results in patients with seminomas are good, whereas results obtained in patients with mediastinal teratocarcinoma, choriocarcinoma, and embryonal carcinoma have been disappointing, with survival times of $< 6$ months (2,5,6).

In connection with this problem, the two patients referred to our Department of Internal Medicine for chemotherapy for a mediastinal germ-cell tumor during the past 2 years were treated with chemotherapy alone. Both patients, one with embryonal carcinoma and the other with seminoma, obtained a complete remission with the combination of vinblastine, bleomycin, and *cis*-dichlorodiammineplatinum(II) (CDDP).

## CASE REPORTS

### Patient 1

Patient 1, a 24-year-old student, gradually developed complaints of tiredness and dyspnea (June 1977). Physical examination revealed dullness of the left hemithorax and a puncture produced hemorrhagic fluid. Thoracotomy showed a tumor arising from the anterior mediastinum and filling the left hemithorax. The histologic diagnosis was embryonal carcinoma locally differentiating towards a yolk sac (endodermal sinus) tumor. The testes were normal on palpation. A lymphangiogram showed normal nodes. The LDH was elevated to 600 units/liter. The α-fetoprotein level was reported to be between 200 and 1000 ng/ml. The β-human chorionic gonadotropin was negative. Polychemotherapy consisting of CDDP (20 mg/m$^2$ for 5 days), vinblastine (0.2 mg/kg of body weight on Days 1 and 2), and bleomycin (30 mg/kg/week) was started. The courses were repeated every 3 weeks, and a total of four courses were given. After the fourth course the tumor seemed smaller. The α-fetoprotein level was 4 ng/ml. In October 1977, a percutaneous puncture produced necrotic material. Thoracotomy showed a large solid tumor arising from the anterior mediastinum and adhering to the pericardium, pulmonary vessels, left main bronchus, and left upper lobe. The mass and the left lung were removed entirely. Histologic examination showed only necrotic tumor tissue. Maintenance therapy consisting of vinblastine (0.2 mg/kg of body weight every 3 weeks) and CDDP (30 mg/m$^2$ [reduced dose because of diminished creatinine clearance] every 6 weeks) was started in November 1977. A complete remission has lasted at least since October 1977 (and probably longer), and the patient is now successfully studying dentistry.

[1]Received July 6, 1979; revised Oct 23, 1979; accepted Oct 29, 1979.

[2]Oncology Unit, Department of Internal Medicine, University of Utrecht, The Netherlands.

[3]*Present address:* Department of Internal Medicine, Free University, Amsterdam, The Netherlands.

[4]We are indebted to Prof. J. van Unnik, Institute for Pathology, University of Utrecht, for revision of the preparations.

*Reprint requests to:* Q. G. C. M. van Hoesel, MD, Oncology Unit, Department of Internal Medicine, University of Utrecht, 3500 CG Utrecht, Catharijnesingel 101, Postbus 16250, The Netherlands.

### Patient 2

Patient 2, a 19-year-old student, was found to have a mediastinal mass during a medical examination for military service in 1974. Histologic examination of the material removed at thoracotomy revealed a seminoma. The testes were normal on palpation. Retroperitoneal exploration was not performed. Postoperatively, the mass was irradiated with a dose of 4000 rads and subsequently disappeared. The retroperitoneal lymph nodes were not irradiated. In August 1975, the patient complained of back pain. A roentgenogram showed a metastasis to the left 11th rib, which was confirmed cytologically. LDH activity was 1300 units/liter. Radiotherapy induced a remission which lasted 5 months.

In January 1976, supraclavicular masses were found on both sides. Histologic examination revealed metastasis of seminoma. Mechlorethamine was started at a dose of 0.1 mg/kg of body weight for 4 days followed by a dose of 0.1 mg/kg every 3 weeks, and remission was again achieved.

In February 1977, his condition worsened. Physical examination showed a superior caval vein syndrome and a mass in the left supraclavicular fossa; the hepar extended 6 cm below the right arcus costae. The lumbar vertebrae were normal. After additional radiotherapy (1500 rads) was applied to the mediastinum, right hilus, and left supraclavicular fossa, the masses disappeared and the superior caval vein syndrome subsided. The chemotherapy was continued in two courses which included actinomycin D (0.5 mg iv for 5 days), vincristine (1 mg on Days 1 and 5), and chlorambucil (8 mg for 14 days).

In June 1977, the patient again complained of back pain. Physical examination showed a radicular syndrome and fractures of the first and second lumbar vertebrae. After radiotherapy (3000 rads), all symptoms disappeared. The patient refused further treatment and had followup until January 1978 when he gradually became dyspneic. Physical examination showed a firm tumor on the left thoracic wall. Cytologic investigation showed an undifferentiated tumor. A roentgenogram showed a mass in the mediastinum, atelectasis of the left lower lobe, and a mass in the right hilus.

In February 1978, the patient was referred to our department for chemotherapy. We decided to give him a course of vinblastine (0.1 mg/kg of body weight on Days 1 and 2) and CDDP (20 mg/m$^2$ for 5 days) to be repeated every 3 weeks. The initial two courses were given without bleomycin because of the possibility of pulmonary fibrosis caused by heavy previous radiation therapy. The LDH activity before the first course was 1500 units/liter; after the second course, the level returned to normal (188 units/liter). The masses disappeared. A complete remission was achieved after the second course and has lasted 26 months. The α-fetoprotein and β-human chorionic gonadotropin subunit have never been elevated. Four courses were given; the third and fourth courses included bleomycin at a weekly dose of 30 mg. Each course resulted in leukopenia and thrombocytopenia, but no septic or hemorrhagic complications occurred. Maintenance therapy, consisting of vinblastine (0.1 mg/kg of body weight every 3 weeks) and CDDP (50 mg/m$^2$ every 6 weeks) was continued until June 1979. At present, the patient is successfully studying law.

## DISCUSSION

In these two patients, the combination of vinblastine, bleomycin, and CDDP induced complete remissions of 32 and 26 months.

The non-seminomatous germ-cell tumors are known to be relatively resistant to therapy (7). After Li et al (8) and Samuels et al (6), Einhorn and Donohue (9,10) achieved major progress in the treatment of non-seminomatous testicular carcinoma.

Surgery has been the first step in the treatment of mediastinal non-seminomatous carcinoma, and was performed where the histologic diagnosis was still unknown. Radiation therapy has not appreciably prolonged survival (2,7). In some patients, however, multiple-drug therapy has been used successfully, with survival times of 30 and 68 months (11). Our first patient had an inoperable tumor diagnosed as embryonal carcinoma. We decided to treat him identically to our patients with testicular germ-cell carcinoma. Because we had frequently found necrotic and fibrotic tumor tissue at retroperitoneal lymph-node dissection after remission induction, we performed a thoracotomy to explore the residual mass and found only necrotic tumor tissue.

Radiotherapy is the treatment of choice for pure testicular seminoma. Reported cure rates reach 95% (5). In this respect, mediastinal seminomas do not seem to differ from the testicular seminomas.

A review of Besznyák et al (12) on mediastinal seminoma indicates a 1-year survival rate of 90% after combined surgical treatment and radiation therapy. Sixteen of 18 patients who were given only radiotherapy lived > 1 year. The favorable influence of radiotherapy has been confirmed by other investigators (7,13–15). The high response rate to radiation therapy permits low doses, so that the heart and lungs are at low-risk with respect to radiation therapy effects. We recommend radiotherapy for mediastinal seminoma after surgery, which is usually required for diagnostic purposes. Radiotherapy for metastatic lesions of testicular seminoma results in a 5-year survival rate of 55% (16). For advanced cases, chemotherapy is the treatment of choice. Alkylating agents are said to give the best results (17,18). Recently, complete remissions obtained with the combination of vinblastine, bleomycin, and CDDP have been reported (9,10). Over

the last 3 years "therapy-resistant" testicular seminoma has been treated like non-seminomatous testicular neoplasms with this drug combination giving excellent results (19). On this basis and the excellent results obtained in patient 2, we propose the use of this drug combination whenever chemotherapy is indicated for widespread metastases of seminoma.

In our opinion, the best way to treat non-seminomatous germ-cell tumors of the mediastinum is by using the combination of vinblastine, bleomycin, and CDDP, thus avoiding high-dose radiotherapy with its harmful effects on the heart and lung, its detrimental effects on the bone marrow reserves, and its high risk of pulmonary fibrosis in conjunction with bleomycin. A recent review describes 19 patients treated with several vinblastine, actinomycin D, and bleomycin protocols resulting in two complete remissions and eight partial remissions, with an overall objective response rate of 53% (20). However, because the incidence of mediastinal germ-cell tumors is so low, it will not be possible to collect sufficiently large series to determine whether it is indeed correct to base the treatment of this kind of neoplasm on the histologic resemblance to testicular carcinoma.

## REFERENCES


1. Prym P. Spontanheilung eines bösartigen, wahrscheinlich chorionepiheliomatösen Gewächses im Hoden. Virchows Arch [Pathol Anat] 265:239–258, 1972.
2. Johnson DE, Laneri JP, Mountain CF, et al. Extragonadal germ cell tumors. Surgery 73:85–90, 1973.
3. Asif S, and Uehling DT. Microscopic tumor foci in testes. J Urol 99:776–779, 1968.
4. Utz DC, and Buscemi MF. Extragonadal testicular tumors. J Urol 105:271–274, 1971.
5. Maier JG, Mittemeyer BT, and Sulak MH. Treatment and prognosis in seminoma of the testis. J Urol 99:72–78, 1968.
6. Samuels ML, Holoye PY, and Johnson DE. Bleomycin combination chemotherapy in the management of testicular neoplasia. Cancer 36:318–326, 1975.
7. Wychulis AR, Payne WS, Clagett OT, et al. Surgical treatment of mediastinal tumors: a 40 year experience. J Thorac Cardiovasc Surg 62:379–392, 1971.
8. Li MC, Whitmore WF, Jr, Golbey R, et al. Effects of combined drug therapy on metastatic cancer of the testis. JAMA 174:1291–1299, 1960.
9. Einhorn LH, and Donohue JP. Improved chemotherapy in disseminated testicular cancer. J Urol 117:65–69, 1977.
10. ——— *cis*-Diamminedichloroplatinum, vinblastine, and bleomycin combination chemotherapy in disseminated testicular cancer. Ann Intern Med 87:293–298, 1977.
11. Walden PAM, Woods RL, Fox B, et al. Primary mediastinal trophoblastic teratomas. Thorax 32:752–758, 1977.
12. Besznyak I, Sebestény M, and Kuchár F. Primary mediastinal seminoma. A case report and review of the literature. J Thorac Cardiovasc Surg 65:930–934, 1973.
13. el-Domeiri AA, Hutter RV, Pool JL, et al. Primary seminoma of the anterior mediastinum. Ann Thorac Surg 6:513–521, 1968.
14. Sterchi M, and Cordell AR. Seminoma of the anterior mediastinum. Ann Thorac Surg 19:371–377, 1975.
15. Polansky SM, Barwick KW, and Ravin CE. Primary mediastinal seminoma. AJR 132:17–21, 1979.
16. Friedman M, and Purkayashtha MC. Recurrent seminoma: management of late metastases, recurrence of a second primary tumor. AJR 83:25–42, 1960.
17. Carter SK, and Wasserman TH. The chemotherapy of urologic cancer. Cancer 36:729–747, 1975.
18. Weitzman SA, and Carey RW. High dose cyclophosphamide as an adjunct to radiotherapy for advanced seminoma. J Urol 117:613–614, 1977.
19. Stoter G, Vendrik CPJ, Struyvenberg A, et al. Combination chemotherapy with *cis*-diammine-dichloro-platinum, vinblastine, and bleomycin in advanced testicular nonseminoma. Lancet 1(8123):941–945, 1979.
20. Reynolds TF, Yagoda A, Golbey R, et al. Chemotherapy of mediastinal germ-cell tumors. Semin Oncol 6:113–115, 1979.

# Treatment of Advanced Ovarian Carcinoma With a Combination of Hexamethylmelamine, Cyclophosphamide, Methotrexate, and 5-Fluorouracil (Hexa-CAF) in Patients With and Without Previous Treatment [1]

Brief Reports and Preliminary Communications

J. P. Neijt,[2,*] A. C. M. van Lindert,[2] C. P. J. Vendrik,[2] K. J. Roozendaal,[3] A. Struyvenberg,[2] and H. M. Pinedo[4]

## SUMMARY

**Forty patients with stage III or IV epithelial ovarian tumors were treated with a combination of hexamethylmelamine, cyclophosphamide, methotrexate, and 5-fluorouracil. Fourteen of 37 evaluable patients had received previous treatment. Three of these 14 previously treated patients responded; no complete remissions were achieved and more severe toxicity was seen compared with the 23 previously untreated patients. In the latter group of patients, seven partial remissions and five complete remissions documented by peritoneoscopy and laparotomy were achieved. It is concluded that the combination of hexamethylmelamine, cyclophosphamide, methotrexate, and 5-fluorouracil is not useful as second-line chemotherapy but can be considered suitable for remission induction in previously untreated patients.**

[Cancer Treat Rep 64:323–326, 1980]

In patients with advanced ovarian carcinoma, Young et al (1) reported improved results with the four-drug combination of hexamethylmelamine, cyclophosphamide, methotrexate, and 5-fluorouracil (Hexa-CAF) compared with melphalan. This report presents our experience with the Hexa-CAF combination in 14 previously treated and 23 previously untreated patients.

## MATERIALS AND METHODS

Only patients with stage III or IV (International Federation of Gynecology and Obstetrics [FIGO] classification) epithelial ovarian tumors were included in this study. The staff of the Pathology Department, the University of Utrecht, reviewed the histology and performed the histologic grading according to Broders (2). Patients were excluded from treatment with Hexa-CAF if they were > 70 years old, and had marked senility, an active second tumor or serious concomitant disease, an unfavorable hematologic status (leukocytes < 3000/mm$^3$ and platelets < 100,000/mm$^3$), or a creatinine clearance < 70 ml/minute. In all patients the extent of disease was determined by surgical exploration supplemented by other procedures such as lymphangiography, barium enema, iv pyelography, and computerized tomographic scanning. During the first laparotomy as much tumor as technically possible was removed, including the uterus, ovaries, and omentum. For patients who initially had had surgical treatment elsewhere, with incomplete tumor reduction, chemotherapy was started and re-exploration was done as soon as clinical examination indicated that the patient seemed to have achieved a remission. When initial tumor-reductive surgery was adequate and only small tumor nodules remained, second-look laparoscopy, followed if necessary by laparotomy, was planned after six courses of chemotherapy. In all patients who had second-look surgery, biopsy of suspect areas and ip lavage were performed. If technically possible, the tumor was further reduced, and the uterus, the remaining ovary or ovaries, and the omentum or other remnants were removed. If < 2 cm of tumor was left after second-look surgery in a responding patient, third-look surgery was done after another six courses of chemotherapy.

The drug regimen was similar to that of Young et al (1). Methotrexate at a dose of 40 mg/m$^2$ and 5-fluorouracil at a dose of 600 mg/m$^2$ were given iv on Days 1 and 8 and cyclophosphamide and hexamethylmelamine were each

[1]Received July 11, 1979; revised Oct 31, 1979; accepted Nov 13, 1979.

[2]Departments of Internal Medicine (*J. P. Neijt, C. P. J. Vendrik,* and *A. Struyvenberg*) and Gynecology (*A. C. M. van Lindert*), University Hospital, Utrecht, The Netherlands.

[3]Department of Gynecology, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands.

[4]Department of Oncology, Free University Hospital, Amsterdam, and the Netherlands Cancer Institute.

**Reprint requests to:* Dr. J. P. Neijt, Department of Internal Medicine, University Hospital, Catharijnesingel 101, 3511 GV Utrecht, The Netherlands.

given orally at a dose of 150 mg/m²/day on Days 1–14. From Day 15 to Day 28 no drugs were given and the cycle was repeated. When there was clear evidence of tumor growth after two or more cycles, the study was discontinued for that patient. If complete remission had been achieved by second- or third-look surgery, the patient continued to receive this treatment for another six cycles. Therapy was discontinued when laparoscopy showed no further evidence of disease after that period. In all cases, the drug-free period was extended after the 12th cycle from Day 15 to Day 42.

Patients were considered evaluable if they had received at least two treatment cycles and were still alive after being in the study > 6 months. Complete remission was defined as the disappearance of all evidence of disease from re-staging by laparoscopy followed by laparotomy. Partial remission was defined as a ≥ 50% reduction of palpable or measurable tumors persisting for at least 1 month. The duration of remission was measured from the first sign of tumor regression.

## RESULTS

Thirty-seven of the 40 patients receiving Hexa-CAF were evaluable. Those patients who were not evaluable included one patient who died 14 days after the start of therapy, one who refused further treatment, and one who was not fully re-staged.

Fourteen of the 37 evaluable patients had previously received extensive radiotherapy or chemotherapy (radiotherapy in one patient, radiotherapy plus chlorambucil in four, and chlorambucil alone in nine). The relevant data and response of the patients with and without previous treatment are shown in table 1. Only three of the 14 previously treated patients achieved a partial remission (4, 11, and 12 months).

Of the 23 patients who had not been previously treated, 15 were classified as having FIGO stage IV disease (nine patients with liver metastases, one with liver metastases and a supraclavicular lymph node, two with affected inguinal lymph nodes, one with an affected supraclavicular lymph node, one with lung metastases, and one with pleural effusion with positive cytology). A partial remission was achieved in six patients with FIGO stage IV disease (measurable abdominal tumor and liver metastases in five patients and inguinal lymph nodes in one). In a patient with small liver metastases and residual tumor < 2 cm prior to chemotherapy, only microscopic amounts of neoplastic tissue were seen at second-look surgery. Unfortunately, third-look surgery showed dis-

TABLE 1.—Characteristics and response in patients with and without previous treatment

| | No. of patients— | | | | |
|---|---|---|---|---|---|
| | Previously treated | With partial response | Previously untreated | With complete response | With partial response |
| Total | 14 | 3 | 23 | 5 (22%) | 7 (30%) |
| Mean age in yrs (range) | 56 (49–68) | | 52 (20–70) | | |
| FIGO stage | | | | | |
| III | 6 | 1 | 8 | 5 | 1 |
| IV | 8 | 2 | 15 | — | 6 |
| Histologic type | | | | | |
| Serous | 5 | 1 | 10 | 3 | 3 |
| Mucinous | — | — | 1 | — | — |
| Undifferentiated | 9 | 2 | 12 | 2 | 4 |
| Histologic grade (Broders, ref 2) | | | | | |
| 1 | — | — | 2 | 2 | — |
| 2 | — | — | 2 | — | — |
| 3 | 2 | — | 9 | 3 | 4 |
| 4 | 12 | 3 | 10 | — | 3 |
| Surgery | | | | | |
| BSO + TAH + omentectomy* | 2 | — | 11 | 4 | 3 |
| BSO + TAH* | 3 | — | 2 | — | 1 |
| Less than above | 9 | 3 | 10 | 1 | 3 |
| Residual tumor prior to chemotherapy | | | | | |
| Microscopic | — | — | 0 | — | — |
| ≤ 2 cm† | — | — | 4 | 2 | 2 |
| > 2 cm | 14 | 3 | 19 | 3 | 5 |

*BSO = bilateral salpingo-oophorectomy; TAH = total abdominal hysterectomy.
†Cross-sectional diameter of residual tumor.

TABLE 2.—Relevant data and findings at 2nd- and 3rd-look surgery for complete responders

| Patient No. | Age/FIGO stage | Cell type/grade | Residual tumor after initial surgery | No. of courses before 2nd-look surgery | Findings at 2nd-look surgery | Findings at 3rd-look surgery (6 courses later) |
|---|---|---|---|---|---|---|
| 1 | 46/III | Undifferentiated/3 | 1–2-mm implants in pelvis and abdominal peritoneum | 6 | Negative | Not done |
| 2 | 34/III | Undifferentiated/3 | 5 × 6-cm fixed plaque at flexura lienalis, 2–3-mm implants in diaphragm | 6 | Partial remission, reductive surgery possible, residual tumor after operation was microscopic | Negative |
| 3 | 70/III | Serous/4 | 2–5-mm implants in pelvis and abdominal cavity | 6 | Negative by laparoscopy | Negative |
| 4 | 32/III | Serous/1 | 3 × 4-cm plaque in recto-uterine pouch | 4 | Biopsies negative (psammoma bodies in ip lavage fluid) | Negative |
| 5 | 26/III | Serous/1 | Fixed large tumor in right and left ovaries, 5 × 5-cm nodules in omentum, diffuse implants in pelvic peritoneum and upper abdomen | 5 | Partial remission, tumor mobile, BSO,* TAH,* omentectomy, and removal of implants, microscopic tumor left | Negative |

**See* table 1.

ease progression. One patient with stage III disease achieved a partial remission. In this case the small tumor volume seen before chemotherapy had become microscopic at second-look surgery (third-look surgery was not done). The median duration of partial remission was 8+ months (range, 3–25). Complete remission was achieved only in patients with stage III disease. Relevant data and findings at second- and third-look surgery are shown in table 2. All of the five patients who achieved complete remission are still alive and disease-free at 7, 13, 22, 25, and 35 months.

## Toxicity

The hematologic toxicity of the regimen is shown in table 3. Comparison of the two groups showed that severe myelosuppression occurred more frequently in previously treated patients. Most of the patients required dose reduction because of myelosuppression. Patients who received Hexa-CAF as initial treatment could be given a mean of 77% of the calculated dose in the first cycle, compared to 63% given to previously treated patients. Mean doses for previously untreated patients decreased in the five successive management cycles: 77%, 74%, 64%, 66%, and 57%. In the previously treated group, no further dose reduction was needed in later cycles (3). There were no drug-related deaths.

Most of the patients experienced gastrointestinal toxicity in the form of nausea and vomiting during the first 14 days of the cycle. Alopecia developed frequently, but hair usually regrew before the sixth treatment cycle.

TABLE 3.—Hematologic toxicity*

| Degree of toxicity | No. of patients— Previously untreated (total, 23 patients) | Previously treated (total, 14 patients) |
|---|---|---|
| None | 2 (8%) | 0 (0) |
| Mild | 6 (26%) | 3 (21%) |
| Moderate | 6 (26%) | 3 (21%) |
| Severe | 9 (39%) | 8 (57%) |

*Mild = wbc count < 3000/mm³ or platelet count < 150,000/mm³; moderate = wbc count < 2000/mm³ or platelet count < 100,000/mm³; severe = wbc count < 1200/mm³ or platelet count < 50,000/mm³.

## DISCUSSION

The induction of a well-documented complete remission in a patient with advanced ovarian cancer is important, because these patients can be expected to have a long disease-free period and may enjoy long-term survival. There is evidence that when maximal surgical resection results in reduction to tumor aggregates < 2 cm, Hexa-CAF chemotherapy, compared to melphalan, is very effective in inducing a complete remission (1). The complete remission rate (22%) in our series is lower than that reported by Young et al (33%) (1), possibly because of the comparatively large number of patients with stage IV disease in our study (15 of 23 patients versus eight of 41 in the study by Young et al) and the small number of patients with residual tumor < 2 cm prior to chemotherapy.

As our data show, the Hexa-CAF regimen is rather toxic. The prolonged treatment is a heavy load and the inevitable dose reductions mean the risk of relapse, especially for those patients who do not achieve a complete remission. It is our opinion, therefore, that powerful but toxic drug combinations such as Hexa-CAF should be reserved only for remission induction. If after 1 year the patient is unlikely to achieve a complete remission, a less toxic combination such as hexamethylmelamine with cyclophosphamide or alkylating monotherapy should be substituted.

Because of the poor response rate and marked hematologic toxicity, Hexa-CAF is not useful as second-line chemotherapy. For this purpose, other new combinations containing *cis*-dichlorodiammineplatinum(II) and Adriamycin seem more promising (4).

## REFERENCES


1. YOUNG RC, CHABNER BA, HUBBARD SP, ET AL. Advanced ovarian adenocarcinoma, a prospective clinical trial of melphalan (L-PAM) versus combination chemotherapy. N Engl J Med 299:1261–1266, 1978.
2. BRODERS AC. Carcinoma: grading and practical application. Arch Pathol 2:376–381, 1926.
3. NEIJT JP, VAN LINDERT ACM, VENDRIK CPJ, ET AL. Hexa-CAF combination chemotherapy and other multiple-drug regimens in advanced ovarian carcinoma: present and future. Neth J Med 22:38–44, 1979.
4. VOGL SE, BERENZWEIG M, KAPLAN BH, ET AL. The CHAD and HAD regimens in advanced ovarian cancer: combination chemotherapy including cyclophosphamide, hexamethylmelamine, Adriamycin, and *cis*-dichlorodiammineplatinum(II). Cancer Treat Rep 63:311–317, 1979.

# Burkitt's Lymphoma: Tumor Lysis Following Malignant Hyperthermia [1]

Brief Reports and Preliminary Communications

Robert P. Suer,* William B. Fisher, Richard W. Pearson, and Douglas A. Triplett [2]

## SUMMARY

**An adult patient with massive abdominal Burkitt's lymphoma developed malignant hyperthermia during exploratory laparotomy. Simultaneously, evidence of tumor lysis appeared. The lethal effect of heat on cancer cells has been amply demonstrated experimentally, and clinical trials of regional and whole-body hyperthermia in various human malignancies have appeared. Lymphomas have not been reported among those tumors studied. Burkitt's lymphoma, a rapidly growing tumor, is likely to respond to hyperthermia, and the evidence reported here should be pursued with controlled clinical trials involving induced hyperthermia in refractory malignant lymphomas.**

[Cancer Treat Rep 64:327–330, 1980]

Neoplastic cells have been shown to be more sensitive to the lethal effects of heat than normal cells (1–3). The original observations were made in patients whose tumors spontaneously regressed or disappeared after febrile illnesses. The lethal effects of heat on cancer cells were later demonstrated in tissue culture and in animal models leading to clinical trials of regional and whole-body hyperthermia in a variety of malignancies (4–7). We wish to report evidence of tumor lysis following an episode of hyperthermia in a patient with Burkitt's lymphoma. This case is unique in that lymphomas have not been reported among those tumors studied with hyperthermia.

## CASE REPORT

Patient 1 is a 27-year-old white male who was admitted to Ball Memorial Hospital on April 18, 1979, with a 1-week history of abdominal pain, distention, nausea, vomiting, and night sweats without fever. Physical examination demonstrated normal vital signs, abdominal distention without hepatosplenomegaly or palpable masses, and no lymph node enlargement. He was slightly tender in both upper quadrants. Admitting laboratory data revealed lymphopenia, hyperuricemia, and mildly elevated liver function tests (*see* table 1). He became increasingly distended over the next 6 days and developed obstructive jaundice. A computerized axial tomographic scan of the abdomen revealed a thick mesenteric "cake" extending from the diaphragm into the pelvis and massive porta hepatis lymph node enlargement. His liver and spleen appeared to be normal. On the tenth hospital day, an exploratory laparotomy was performed. The tumor was an adherent mass occupying the entire peritoneal cavity. It was not possible to identify the small bowel or colon since these were entirely encased in tumor. The liver edge was visible and appeared grossly normal. The tumor itself was extremely friable and immobile. Superficial tumor, liver, and bilateral bone marrow biopsies were obtained. The porta hepatis and para-aortic areas could not be visualized. The biopsy was consistent with Burkitt's lymphoma (figs 1 and 2). The liver biopsy showed cholestasis, and the bone marrow biopsies were normal. During and after surgery, the patient experienced features of malignant hyperthermia: strong, sustained muscle contractions after succinylcholine, persistent tachycardia, profound respiratory and metabolic acidosis, hyperkalemia accompanied by characteristic electrocardiographic changes, ventricular tachycardia, and hyperthermia. His temperature was not monitored during surgery but was recorded at 103.5°F after surgery. He responded to iv procainamide and external cooling. At no time was he oliguric or hypotensive. Over the next 72 hours his obstructive liver function pattern improved dramatically accompanied by a marked elevation in serum uric acid (*see* table 1). He was treated with cyclophosphamide and intrathecal cytosine arabinoside beginning on Day 14 and was transferred to the National Cancer Institute (Bethesda, MD). A muscle biopsy was later ob-

[1]Received July 20, 1979; revised Nov 13, 1979; accepted Nov 26, 1979.

[2]Departments of Medicine (*R. P. Suer*), Medical Education (*W. B. Fisher*), and Pathology (*R. W. Pearson* and *D. A. Triplett*), Ball Memorial Hospital, Muncie, IN.

**Reprint requests to:* Dr. Robert P. Suer, Department of Medicine, Ball Memorial Hospital, 2401 University Ave, Muncie, IN 47303.

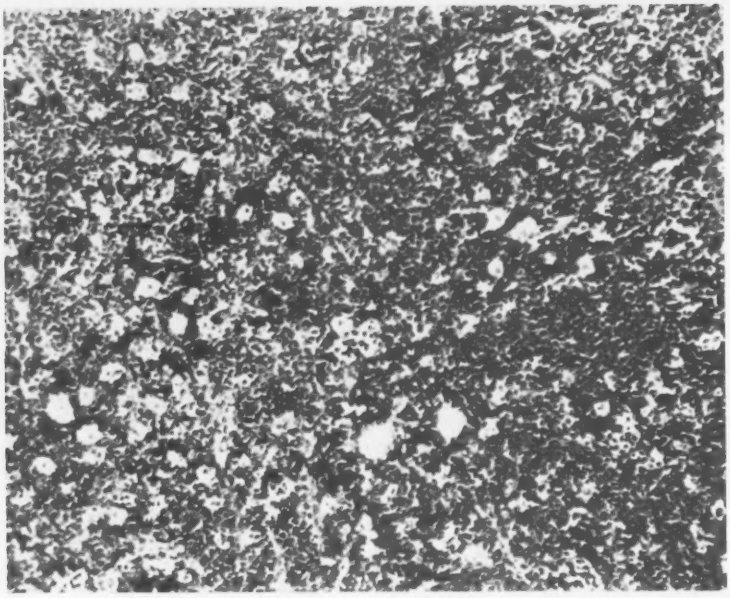


FIGURE 1.—"Starry-sky" pattern created by the presence of large, nonneoplastic foamy histiocytes amidst a background of undifferentiated lymphoid cells (hematoxylin and eosin stain, × 100) (ref 8).

TABLE 1.—Liver function tests in patient 1

| Serum | Serum level on hospital Day*— | | | |
|---|---|---|---|---|
| | 1 | 9 | 11 | 13 |
| Alkaline phosphatase (milliunits/ml) | 240 | 1310 | 1080 | 615 |
| LDH (milliunits/ml) | 355 | 550 | 560 | 491 |
| SGOT (milliunits/ml) | 232 | 560 | 400 | 105 |
| Uric acid (mg/dl) | 11.7 | 7.4 | 17.0 | 7.4 |
| Total bilirubin (mg/dl) | 0.8 | 7.4 | 2.3 | 1.7 |
| CA++ (mg/dl) | 9.0 | 9.2 | 7.5 | 8.2 |
| Phosphorus (mg/dl) | 3.3 | 5.1 | 4.9 | 2.2 |
| Potassium (mEq/liter) | 3.7 | 4.5 | 7.1 | 3.4 |
| BUN (mg/dl) | 12 | 13 | 26 | 8 |

* Surgery was performed on Day 10.

tained and analyzed in vitro.[3] Two indications of abnormal sarcoplasmic reticulum were obtained: abnormal caffeine hypersensitivity and delayed relaxation of muscle tension. Both of these manifestations are typical of malignant hyperthermia.

[3]Wood DS. H. Houston Merritt Clinical Research Center. College of Physicians and Surgeons of Columbia University, New York, NY. Personal communication.

## DISCUSSION

Patients with Burkitt's lymphoma and large tumors frequently undergo serious metabolic derangements following chemotherapy as a result of extensive tumor cell lysis (9). This is associated with hyperkalemia, hypocalcemia, hyperphosphatemia, azotemia, and hyperuricemia. A similar metabolic picture is associated with malignant hyperthermia as a result of loss of muscle cell membrane integrity (10). Malignant hyperthermia could

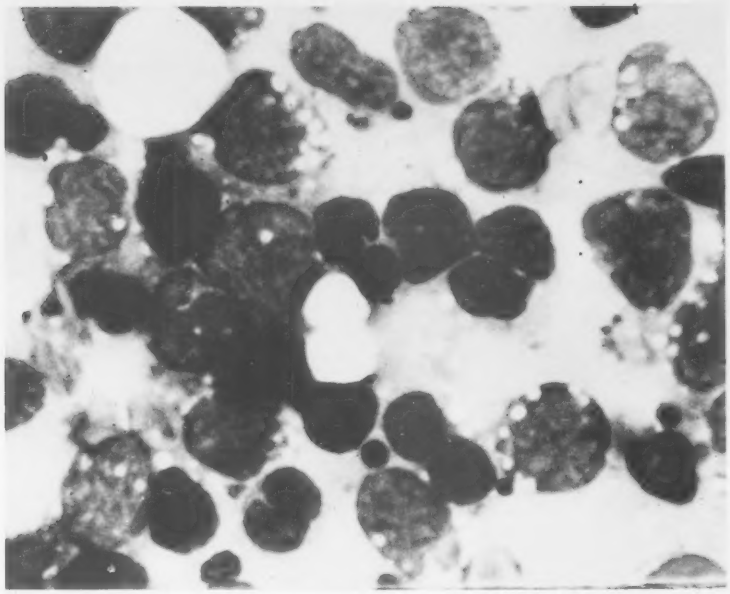


FIGURE 2.—Imprint showing undifferentiated lymphoid cells with 2–5 faint nucleoli, basophilic cytoplasm, and multiple cytoplasmic vacuoles. Tumor cells were methyl green-pyronine positive, oil red-O positive, and PAS negative (Wright's stain, × 1000) (ref 8).

not account for the dramatic relief of our patient's obstructive jaundice or the marked hyperuricemia occurring immediately following surgery. It is highly doubtful that surgical manipulation resulted in the above picture since the biliary system was inaccessible. Though it cannot be definitely concluded, we suggest that the most logical interpretation of the clinical and laboratory evidence of relief of obstructive jaundice and transient hyperuricemia is hyperthermia-induced tumor regression.

Our case is the third report of malignant hyperthermia in Burkitt's lymphoma (11). The rarity of these disorders suggests a possible association. Viral particles have been identified in muscle cells of patients with an apparently acquired form of malignant hyperthermia and may be of etiologic importance (12). It is an intriguing possibility that a common viral etiology induces both conditions.

Lymphomas have not been mentioned in reports on the use of hyperthermia in malignancy. This is probably a consequence of the success of the traditional chemotherapeutic regimens in treating these disorders. Tumors with high growth fractions are more sensitive to heat than those with lower values (1). Burkitt's lymphoma, a rapidly growing tumor, is a likely candidate to respond to hyperthermia. We believe that this evidence should be pursued with controlled clinical trials involving induced hyperthermia.

## REFERENCES


1. CAVALIERE R, CIOCATTO EC, GIOVANELLA BC, ET AL. Selective heat sensitivity of cancer cells. Cancer 20:1351–1381, 1967.
2. PALZER RJ, and HEIDELBERGER C. Studies on the quantitative biology of hyperthermic killing of HeLa cells. Cancer Res 33:415–421, 1973.
3. GIOVANELLA BC, STEHLIN JS, JR, and MORGAN AC. Selective lethal effect of supranormal temperatures on human neoplastic cells. Cancer Res 36:3944–3950, 1976.
4. WARREN SL. Preliminary study of the effect of artificial fever upon hopeless tumor cases. Am J Roentgenol Rad Ther 33:75–87, 1935.
5. GIOVANELLA BC, LOHMAN WA, and HEIDELBERGER C. Effects of elevated temperatures and drugs on the viability of L1210 leukemic cells. Cancer Res 30:1623–1631, 1970.
6. PETTIGREW RT, GALF JM, LUDGATE CM, ET AL. Clinical effects of whole-body hyperthermia in advanced malignancy. Br Med J 4:679–682, 1974.
7. BULL JM, LEES D, SCHUETTE W, ET AL. Whole body hyperthermia: a phase-1 trial of a potential adjuvant to chemotherapy. Ann Intern Med 90:317–323, 1979.
8. BERARD C, O'CONOR GT, THOMAS LB, ET AL. Histopathologic definition of Burkitt's tumor. Bull WHO 40:601–607, 1969.

9. Arseneau JC, Canellos GP, Banks PM, et al. American Burkitt's lymphoma: a clinicopathologic study of 30 cases. I. Clinical factors relating to prolonged survival. Am J Med 58:314–321, 1975.
10. Britt BA. Malignant hyperthermia: a pharmacogenetic disease of skeletal and cardiac muscle. N Engl J Med 290:1140–1142, 1974.
11. Tsueda K, Dubick MN, Wright BD, et al. Intraoperative hyperthermia crisis in two children with undifferentiated lymphoma. Anesth Analg 57:511–514, 1978.
12. Shiller HH, and Mair WGP. Ultrastructural changes of muscle in malignant hyperthermia. J Neurol Sci 21:93–100, 1974.

# Vinblastine, Intermittent Bleomycin, and Single-Dose *cis*-Dichlorodiammineplatinum(II) in the Management of Stage III Testicular Cancer[1,2]

Brief Reports and Preliminary Communications

George J. Bosl,[3,4] Roger Kwong,[3,5] Paul H. Lange,[3] Elwin E. Fraley,[3] and B. J. Kennedy[3,*]

## SUMMARY

**A program for treating advanced testicular cancer with vinblastine, bleomycin, and a single dose of *cis*-dichlorodiammineplatinum(II) was used in an attempt to reduce the toxicity of a 5-day regimen while maintaining a high response rate. In ten previously untreated and six previously treated patients, the complete remission rate was 25%. Toxic reactions were frequent, and there was one drug-related death. This single-dose *cis*-dichlorodiammineplatinum(II) program is ineffective in the treatment of testicular cancer. Other means for reducing toxicity will have to be considered.**

[Cancer Treat Rep 64:331–334, 1980]

Chemotherapy for stage III testicular cancer has improved since 1970. Vinblastine with bleomycin by either intermittent injection or infusion induces complete or partial remissions in a substantial percentage of patients, offering hope of cure to patients who previously had a bleak prognosis. Phase I studies of *cis*-dichlorodiammineplatinum(II) (cisplatin) by Higby et al (1) demonstrated the considerable activity of this agent against testicular carcinoma. Subsequently, a combination of vinblastine, bleomycin, and cisplatin was used in a 5-day program with further enhancement of antitumor effect; complete remissions were seen in 15 of 20 (75%) evaluable patients (2). However, significant toxic effects occurred: nausea, vomiting, alopecia, stomatitis, leukopenia with septicemia, and nephrotoxicity. The last toxic effect limits the dose of cisplatin, and gastrointestinal reactions can make patients less cooperative. Nephrotoxicity can be ameliorated by vigorous hydration with or without mannitol diuresis (3). We decided to decrease the dose of vinblastine and to change the dose of cisplatin to a smaller, single, iv dose with forced diuresis so that (*a*) the hematologic toxicity would be reduced, (*b*) the duration of gastrointestinal reactions and of hospitalization would be shortened, and (*c*) the nephrotoxicity would be avoided.

## MATERIALS AND METHODS

Between May 1976 and April 1977, 16 patients with stage III testicular cancer were treated with vinblastine, bleomycin, and cisplatin; ten of these patients had not been previously treated and the other six had received extensive chemotherapy, including vinblastine and bleomycin. The regimen consisted of vinblastine at a dose of 0.15 mg/kg iv on Days 1 and 2; bleomycin at a dose of 30 mg iv on Days 2, 7, and 14; and cisplatin at a dose of 60 mg/m$^2$ on Day 2. For 12 hours before the cisplatin administration, the patient was given parenteral fluids to ensure a urine flow of $\geq$ 100 ml/hour. Immediately before cisplatin was given, 12.5 g of mannitol was administered rapidly iv. Cisplatin was then given iv over 20 minutes. Antiemetics were given as necessary, and fluid lost by vomiting was replaced. Cycles were repeated at 3- to 4-week intervals. All but one of these patients received at least two cycles of chemotherapy and most patients received three or more. The exception, patient 1, died of progressive disease within 1 week after treatment started. Maintenance therapy for patients achieving complete remission consisted of vinblastine at a dose of 0.2 mg/kg iv every 3–4 weeks for 2 years, modified as necessary because of toxic reactions.

Complete remission was defined as the regression of all measurable disease with normal levels of human cho-

[1]Received July 25, 1979; accepted Dec 4, 1979.

[2]Supported by grants from the American Cancer Society, the Minnesota Medical Foundation, and the Masonic Hospital Fund. Dr. Bosl is a recipient of the American Cancer Society Junior Clinical Faculty Fellowship.

[3]Section of Medical Oncology, Department of Medicine (*G. L. Bosl, R. Kwong,* and *B. J. Kennedy*), and Department of Urologic Surgery (*P. H. Lange* and *E. E. Fraley*), University of Minnesota College of Health Sciences, Minneapolis.

[4]*Present address:* Solid Tumor Service, Department of Medicine, Sloan-Kettering Memorial Cancer Center, New York, NY.

[5]*Present address:* The Gunderson Clinic, LaCrosse, WI.

**Reprint requests to:* Dr. B. J. Kennedy, Section of Medical Oncology, Department of Medicine, Box 286 Mayo, University of Minnesota, Minneapolis, MN 55455.

rionic gonadotropin (HCG) and alpha-fetoprotein (AFP). A partial remission was defined as a 50% reduction in the size of all measurable lesions or in the size of retroperitoneal metastases as estimated by abdominal echogram or computerized abdominal tomogram. Patients who had only a slight response, no response, or disease progression, were classified as having no response.

Pulmonary and renal functions were monitored throughout induction therapy. The diffusion capacity of carbon monoxide was measured frequently, and the 24-hour creatinine clearance was determined before each cycle. Survival was calculated from the date chemotherapy was started.

## RESULTS

### Clinical and Laboratory Data

The clinical histories of these 16 patients are summarized in tables 1 and 2. The median age was 24.5 years (range, 18–46), and all patients had either recurrent stage II or III disease. Lung metastases were noted in ten patients, retroperitoneal metastases were noted in 13, supraclavicular lymph node metastases were noted in five, liver metastases were noted in three, and brain, mediastinal, or bone metastases were each noted in one. Prior to treatment, the AFP was elevated in seven patients, the HCG was elevated in 13 patients, and both markers were elevated in five patients.

All histologic types of tumor were represented. Two patients had pure seminoma and elevated levels of HCG. Seven patients had pure embryonal carcinoma and another patient had embryonal carcinoma with seminoma. Teratocarcinoma was seen in six patients; in one of these patients the tumor also contained seminoma and in another the tumor contained elements of choriocarcinoma.

### Remission Induction

Complete remission was achieved in three of ten previously untreated patients but in only one of the six patients who had received previous chemotherapy. Partial remission was achieved in two of ten previously untreated patients and in one of six previously treated patients. Thus, the total response rate was 50% for previously untreated patients and 33% for previously treated patients.

### Survival

Five patients are still alive with a median survival time of > 25 months (range, > 24 to > 28). Three of these patients achieved complete remission and maintenance therapy has been discontinued. The other two patients

TABLE 1.—Clinical histories of patients who had not received previous chemotherapy

| Patient No. | Tumor type (stage)* | Sites of metastases | AFP/ HCG† | No. of drug cycles | Response | Status (survival in mos) |
|---|---|---|---|---|---|---|
| 1 | TC(III) | Lung, retroperitoneum | –/+ | 1 | None | Died on Day 4 from cancer |
| 2 | E(R-II) | Retroperitoneum | +/– | 4 | Complete | Alive (> 24); maintenance therapy discontinued (> 12) |
| 3 | S(III) | Liver, lung, retroperitoneum | –/+ | 3 | Minimal | Alive (> 25) after partial response with cyclophosphamide |
| 4 | TC(III) | Liver, lung, retroperitoneum | –/+ | 4 | Partial | Dead (12) |
| 5 | E(III) | Lung, retroperitoneum, supraclavicular node | +/+ | 5 | Minimal | Dead (8) |
| 6 | E(III) | Lung | –/+ | 2 | Complete | Alive (> 25); maintenance therapy discontinued (> 14) |
| 7 | S(III) | Lung, brain, retroperitoneum | –/+ | 5 | None | Dead (10) |
| 8 | S + TC(III) | Retroperitoneum, supraclavicular node | –/– | 5 | Minimal | Alive (> 27); teratoma removed at operation |
| 9 | E(III) | Lung, retroperitoneum, supraclavicular node | +/+ | 7 | Partial | Dead (14); poor compliance with program |
| 10 | TC +C(III) | Mediastinum | –/+ | 4 | Complete | Alive (> 28); maintenance therapy discontinued (> 14) |

*TC = teratocarcinoma, E = embryonal carcinoma, S = seminoma, C = choriocarcinoma; R-II = recurrent retroperitoneal disease, III = metastases outside retroperitoneum.

†+ = elevated at start of treatment; – = normal at start of treatment.

TABLE 2.—Clinical histories of patients who had received previous chemotherapy

| Patient No. | Tumor type (stage)* | Previous therapy† | Sites of metastases | AFP/ HCG* | No. of drug cycles | Response | Status (survival in mos) |
|---|---|---|---|---|---|---|---|
| 11 | E(III) | Vinblastine, bleomycin | Retroperitoneum, supraclavicular node | +/– | 4 | Complete | Dead (14); suicide |
| 12 | TC(III) | T, vinblastine, bleomycin, Adriamycin, cisplatin | Lung, retroperitoneum | +/+ | 2 | Minimal | Dead (4) |
| 13 | E(R-II) | T, mithramycin, Adriamycin, vinblastine, bleomycin | Lung, liver, retroperitoneum | –/+ | 4 | Progression | Dead (3); died of bleomycin toxicity; autopsy showed residual tumor |
| 14 | TC(R-II) | T, vinblastine, bleomycin | Retroperitoneum | +/+ | 4 | Progression | Dead (12) |
| 15 | E(III) | Mithramycin, vinblastine, bleomycin | Lung, supraclavicular node | +/+ | 4 | Partial | Dead (16) |
| 16 | S + E(R-II) | Mithramycin, vinblastine, bleomycin | Bone, retroperitoneum | –/+ | 3 | Progression | Dead (4) |

*See footnotes in table 1 for explanations.
†T = triple therapy: chlorambucil + methotrexate + actinomycin D.

achieved partial remission. One patient (No. 3) received three cycles of chemotherapy for classical seminoma. The serum HCG level rapidly returned to normal, but there were residual lung metastases. Two cycles of a 5-day cisplatin program were given without improvement. The patient began receiving daily doses of cyclophosphamide and has been healthy with stable lung nodules for 2 years. Another patient (No. 8) appeared to have only a slight response to the chemotherapy, and an exploratory operation was performed. However, the resected tumor contained only adult (mature) teratoma. The patient continues in good health.

Another patient (No. 11) achieved complete remission which lasted 14 months. He committed suicide while despondent after being deserted by his wife. No autopsy was performed.

The other ten patients did not respond to chemotherapy and died within 12 months from the start of treatment. In one patient (No. 13), bleomycin toxicity was the immediate cause of death.

## Toxicity

Toxicity could be assessed in only 15 patients because one patient (No. 1) died 4 days after initiation of treatment. Nausea, vomiting, and alopecia were seen in 14 patients. Less common toxic reactions included myalgias (seven patients), constipation (six), drug fever (six), cutaneous erythema or hyperpigmentation (five), mucositis (five), symptomatic peripheral neuropathy (two), and hearing impairment (one).

Serial blood cell counts were available for 13 patients. Leukopenia occurred in all patients, with a median wbc count nadir of 1450/mm$^3$ (range, 900–2900). Thrombocytopenia (< 150,000/mm$^3$) occurred in 11 patients, with a median platelet count nadir of 139,000/mm$^3$ (range, 38,000–180,000). No patients were hospitalized for leukopenia with septicemia or for hemorrhage and thrombocytopenia.

A decrease in the diffusion capacity of carbon monoxide of > 30% of the pretreatment value was noted in two patients. Another patient, who had been extensively treated with bleomycin earlier, died of bleomycin pneumonitis. An autopsy revealed interstitial pneumonitis and residual testicular carcinoma.

Progressive decrease in the creatinine clearance was seen in six patients, two of whom had a decrease to < 50 ml/minute.

## DISCUSSION

A modified program of vinblastine, bleomycin, and cisplatin, used in an attempt to reduce the adverse effects of vinblastine and cisplatin, was ineffective. The complete remission rate of 25% only approaches that of vinblastine and intermittent bleomycin (3), and toxicity was not reduced. There was one drug-related death, and moderately severe leukopenia and thrombocytopenia and severe gastrointestinal symptoms were common.

The failure of this program may be due either to the smaller dose of cisplatin or to the shorter course. Although the dose of vinblastine that was used is less than that originally used by Einhorn et al (2), this dose is effective with less hematologic toxicity, provided the higher dose of cisplatin in the 5-day program is used (4). The VAB III program (5) uses a dose of cisplatin similar

to that of Einhorn et al (2), but the cisplatin is given as a single rapid infusion during mannitol diuresis and hydration. Perhaps, if cisplatin had been given in doses of 100–120 mg/m$^2$ in our program, a better response rate would have been achieved.

Other means are being considered to decrease drug toxicity. One possibility is a 5-day continuous infusion of cisplatin (6). Another possibility is timing the administration of cisplatin for the most sensitive stage of the cell cycle, as suggested by chronopharmacologic studies (7,8). In addition, the optimum duration of maintenance therapy is unknown; shorter schedules may be possible (9).

We conclude that the modified regimen of vinblastine, bleomycin, and cisplatin is not sufficient for the treatment of advanced testicular cancer. Regimens including high-dose cisplatin are clearly the most effective programs now available for this disease, and other methods for reducing toxicity are needed.

## REFERENCES


1. Higby DJ, Wallace HJ, Albert DJ, et al. Diaminodichloroplatinum: a phase I study showing responses in testicular and other tumors. Cancer 33:1219–1225, 1974.
2. Einhorn LH, Furnas E, and Powell N. Combination chemotherapy of differentiated testicular carcinoma with *cis*-platinum diamminodichloride (CPDD), vinblastine (vlb), and bleomycin (bleo). Proc Am Assoc Cancer Res and ASCO 17:240, 1976.
3. Samuels ML, Holoye PY, and Johnson DE. Bleomycin combination chemotherapy in the management of testicular neoplasia. Cancer 36:318–326, 1975.
4. Einhorn LH, and Williams SO. *cis*-Diamminodichloroplatinum(II) plus vinblastine plus bleomycin plus or minus Adriamycin in disseminated testicular cancer. Proc Am Assoc Cancer Res and ASCO 20:297, 1979.
5. Cvitkovic E, Hayes D, and Golbey R. Primary combination chemotherapy (VAB III) for metastatic or unresectable germ cell tumors. Proc Am Assoc Cancer Res and ASCO 17:296, 1976.
6. Salem P, Hall SW, Benjamin RS, et al. Clinical phase I–II study of *cis*-dichlorodiammineplatinum(II) given by continuous iv infusion. Cancer Treat Rep 62:1553–1555, 1978.
7. Hrushesky W, Levi F, Halberg F, et al. *Cis*-platinum chronotolerance. (Abstr) *In* Proceedings of the Conference on Combined Modality Chemotherapy and Radiotherapy, 1978. In press.
8. Levi F, Hrushesky W, Halberg F, et al. Murine chronotolerance to *cis*-platinum (DDP) and its further optimization by hydration. Proc Am Assoc Cancer Res and ASCO 20:274, 1979.
9. Bosl G, Lange P, Fraley E, et al. Is maintenance therapy necessary following complete remission of stage III testis cancer? Proc Am Assoc Cancer Res and ASCO 20:385,1979.

# Combination Chemotherapy With VM-26, Adriamycin, Bleomycin, and Prednisone as a Secondary Treatment of Malignant Lymphoma[1,2]

Brief Reports and Preliminary Communications

A. Goldhirsch,* M. Pirovino, R. W. Sonntag, L. Tschopp, H. J. Ryssel, and K. W. Brunner[3,4]

## SUMMARY

**Thirty-eight pretreated patients with Hodgkin's disease (HD) and malignant non-Hodgkin's lymphoma were given combination chemotherapy with VM-26, Adriamycin, bleomycin, and prednisone. Four of 15 evaluable patients with HD achieved a partial remission (PR), with a median duration of 8 months. Of 12 patients with diffuse poorly differentiated lymphocytic lymphoma, one achieved a complete remission (30+ months) and five achieved a PR (median, 6 months). One of three patients with histiocytic lymphoma had a PR of 1.5 months. There was one drug-related death. Five patients developed life-threatening hematologic toxicity. Two HD responders died of acute nonlymphocytic leukemia.**

[Cancer Treat Rep 64:335–337, 1980]

Adriamycin (ADM), bleomycin (BLM), and prednisone (1–4), as well as epipodophyllotoxin derivatives (5–7), are effective agents in the treatment of malignant lymphomas. VM-26 is a semisynthetic derivative of podophyllotoxin and has a cytotoxic action by blocking the cell cycle in the premitotic stage (8). Clinical studies with VM-26 have shown remission rates of ≤ 36% in Hodgkin's disease (HD), 42% in reticulum cell sarcoma (9), and 45% in lymphosarcoma (7). In our own series (5) we observed a remission rate of 69% in HD; two of 16 patients achieved complete remission (CR).

## PATIENTS AND METHODS

From September 1972 to April 1977, 38 patients with HD and non-Hodgkin's lymphoma (NHL) were treated every 21 days with the combination of VM-26 (40 mg/m$^2$ iv on Days 1 and 2), ADM (40 mg/m$^2$ iv on Day 2), BLM (15 mg iv on Days 1 and 2), and prednisone (40 mg/m$^2$ orally on Days 1–8). Eligible patients had progressive and measurable stage III or IV previously treated HD or treatment-resistant NHL. Eastern Cooperative Oncology Group criteria for response were used.

Eighteen patients with HD and 20 patients with NHL were entered in the study. All patients were pretreated, with 25 of the 38 patients having had at least two different chemotherapeutic combinations, with or without radiotherapy. The mean age of patients with NHL was 52.4 years (range, 21–74) and the mean age of patients with HD was 36.5 years (range, 23–69). There were 11 males and seven females with HD and 11 males and nine females with NHL. All but two patients had stage IV disease. One patient with HD and one with diffuse poorly differentiated lymphocytic lymphoma had stage III disease. All of the 18 patients with HD and 13 of the 20 patients with NHL had B symptoms.

## RESULTS

*Hodgkin's disease.*—Of 18 patients, 15 were evaluable for response, two were unevaluable because of non-treatment-related early death, and one was unevaluable because of protocol violation. No CRs were obtained. Four patients had partial remissions (PRs), eight had no change, and three had progressive disease (table 1). Of six patients with nodular sclerosis, two achieved PRs (5 and 6+ months), three had no change (2, 3, and 6 months), and one had progressive disease. Of five patients with mixed cellularity, one achieved PR (14 months) and four had no change (3, 4, 5, and 8 months). Of four patients with lymphocyte depletion, one achieved PR (8 months), one had no change (1 month), and two had progressive disease.

*Malignant non-Hodgkin's lymphoma.*—Of 20 patients, 17 were evaluable for response and three were unevaluable because of early death. Two patients with nodular poorly differentiated lymphocytic lymphoma had

[1]Received July 31, 1979; revised Oct 30, 1979; accepted Nov 7, 1979.

[2]VM-26 was supplied in part by Dr. M. Jokus of Sandoz AG, Basel, Switzerland.

[3]Division of Oncology, University Hospital, Inselspital Bern, Switzerland.

[4]We gratefully acknowledge the secretarial assistance of Edith Eicher and Bernadette Drittenbass.

*Reprint requests to: A. Goldhirsch, MD, Division of Oncology, University Hospital, Inselspital, 3010 Bern, Switzerland.

TABLE 1.—Relation of prior treatment to results

| Prior treatment | HD | | | | | NHL | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. of patients | | | | | No. of patients | | | | |
| | Total | CR + PR | No change | Progressive disease | Pretreatment study drugs* | Total | CR + PR | No change | Progressive disease | Pretreatment study drugs* |
| Radiotherapy alone | — | — | — | — | — | 1 | — | — | 1 | — |
| No. of prior chemotherapy combinations (± radiotherapy) | | | | | | | | | | |
| 1 | 1 | 1 | — | — | — | 9 | 6 | 1 | 2 | — |
| 2 | 5 | 1 | 3 | (1) | BLM/VP-16-213 | 3 | 2 | 1 | — | — |
| 3 | 6 | 2(1) | 2 | 2 | BLM | 3 | 2(1) | — | 1† | BLM/VM-26 |
| ≥4 | 3 | — | 3(3) | — | ADM/BLM (2) VP-16-213 (1) | | | | | ADM/BLM/VM-26† |
| | — | — | — | — | | 1 | — | — | 1 | |
| Totals | 15 | 4 | 8 | 3 | (5) | 17 | 10 | 2 | 5 | (2) |

*Nos. in parentheses = total No. of pretreated patients.
†Patient was treated with pretreatment study drugs at different times.

TABLE 2.—Hematologic toxicity

| Myelosuppression of— | No. of patients: Eastern Cooperative Oncology Group scale 1 | 2 | 3 | 4 | 5 | Total No. of evaluable patients |
|---|---|---|---|---|---|---|
| Hemoglobin | 11 | 7 | 6 | — | — | 30 |
| Wbcs | 5 | 9 | 10 | 5 | — | 36 |
| Platelets | 8 | 9 | 4 | 3 | 1 | 35 |

no change (2 and 8 months). Of 12 patients with diffuse poorly differentiated lymphocytic lymphoma, one patient achieved CR (30+ months), five achieved PRs (median, 6 months), two had no change (1 and 2 months), and four had progressive disease. Of three patients with diffuse histiocytic lymphoma, one achieved PR (2 months), one had no change (1 month), and one had progressive disease (table 1).

*Toxicity (Eastern Cooperative Oncology Group Criteria).*—Hematologic toxicity is summarized in table 2. Six patients required erythrocyte substitution. A 70-year-old patient died of massive gastrointestinal bleeding due to severe drug-related thrombocytopenia. Severe stomatitis was observed in one patient and moderate stomatitis was observed in 13 patients. Nausea and vomiting was severe in three patients. Sixteen patients experienced controllable vomiting or nausea alone. Diarrhea without dehydration was observed in two patients. Mild to severe alopecia occurred in 22 patients; 16 patients were not evaluable for this toxic effect because of preexistent hair loss. Skin toxicity was observed in two patients (photosensitivity in one and thickening of pressure points on the hands in the other). Two patients developed herpes zoster infection while receiving treatment. In one patient, progression of congestive heart failure occurred. Two heavily pretreated patients with HD developed acute nonlymphocytic leukemia.

## DISCUSSION

The good response rates of VM-26 as a single agent in the treatment of malignant lymphomas (5) encouraged us to include this drug in a combination regimen with ADM, BLM, and prednisone. Our results in these heavily pretreated patients seem promising. Good palliation was afforded, even to patients with stable disease. The drug combination was fairly well-tolerated: one toxic death due to bleeding occurred in a patient with massive bone marrow involvement. The two acute leukemias that were observed developed after a total of 7 and 8 years of treatment with cytostatic agents and radiotherapy (10). ADM and BLM have become part of primary treatment regimens in HD (3) and malignant NHLs (4,11). Our results support the impression of other investigators (7) that VM-26 is an active drug in the treatment of these diseases, but its exact place within the chemotherapeutic armamentarium has yet to be defined. Recent data suggest a possible synergism between epipodophyllotoxin derivatives and *cis*-dichlorodiammineplatinum(II) in other malignancies. We feel, therefore, that this combination, or a combination of VM-26 with other active agents, warrants further clinical investigation.

## REFERENCES


1. BLUM RH, and CARTER SK. Adriamycin: a new anticancer drug with significant clinical activity. Ann Intern Med 80:249–259, 1974.
2. BONADONNA G, DE LENA M, MONFARDINI S, ET AL. Combination usage of Adriamycin (NSC-123127) in malignant lymphomas. Cancer Chemother Rep Part 3, vol 6(2):381–388, 1975.
3. BONADONNA G, USLENGHI C, and ZUCALI R. Recent trends in the medical treatment of Hodgkin's disease. Eur J Cancer 11:251–266, 1975.
4. MONFARDINI S, TANCINI G, DELENA M, ET AL. Cyclophosphamide, vincristine, and prednisone (CVP) versus Adriamycin, bleomycin, and prednisone (ABP) in stage IV non-Hodgkin's lymphomas. Med Pediatr Oncol 3:67–74, 1977.
5. SONNTAG RW, SENN HJ, NAGEL G, ET AL. Experience with 4'-demethylepipodo-phyllotoxin 9-(4,6-*O*-2-thenylidene-beta-D-glucopyranoside); VM-26; NSC-122819 in the treatment of malignant lymphosis. Eur J Cancer 10:93–98, 1974.
6. CHIUTEN DF, BENNETT JM, CREECH RH, ET AL. VM-26, a new anticancer drug with effectiveness in malignant lymphoma: an Eastern Cooperative Oncology Group Study (EST 1474). Cancer Treat Rep 63:7–11, 1979.
7. RADICE PA, BUNN PA, JR, and IHDE DC. Therapeutic trials with VP-16-213 and VM-26: active agents in small cell lung cancer, non-Hodgkin's lymphomas, and other malignancies. Cancer Treat Rep 63:1231–1239, 1979.
8. STAEHELIN H. VM-26—A podophyllum compound with a new mechanism of action. Eur J Cancer 6:303–311, 1970.
9. ROZENCWEIG M, VON HOFF DD, HENNEY JE, ET AL. VM-26 and VP-16-213: a comparative analysis. Cancer 40:334–342, 1977.
10. CAVALLI F, SONNTAG RW, and BRUNNER KW. Acute leukemia in Hodgkin's disease. Scand J Haematol 19:319–320, 1977.
11. SCHEIN PS, DEVITA VT, JR, HUBBARD S, ET AL. Bleomycin, Adriamycin, cyclophosphamide, vincristine, and prednisone (BACOP) combination chemotherapy in the treatment of advanced diffuse histiocytic lymphoma. Ann Intern Med 85:417–422, 1976.

# Phase II Evaluation of *N*-(Phosphonacetyl)-L-aspartic Acid (PALA) in Metastatic Adenocarcinoma of the Colon or Rectum [1]

Brief Reports and Preliminary Communications

David A. Van Echo,* Charles H. Diggs, Mary Scoltock, and Peter H. Wiernik [2]

## SUMMARY

***N*-(Phosphonacetyl)-L-aspartic acid (PALA) was administered at a dose of 1.5 g/m²/day × 5 days in a phase II evaluation of 21 patients with metastatic adenocarcinoma of the colon or rectum. Courses were repeated every 3 weeks. No responses were seen, but 11 patients had disease stabilization. PALA toxicity was marked, causing discontinuation of therapy in two patients. Major toxicity was cutaneous (90%), mucosal (35%), and gastrointestinal (diarrhea, at times bloody) (50%). There was no evidence of myelosuppression or CNS toxicity. PALA appears inactive in patients with metastatic colorectal carcinoma in the dose and schedule used in this study.**

[Cancer Treat Rep 64:339–342, 1980]

Patients with metastatic adenocarcinoma of the colon or rectum have poor response rates to all known chemotherapeutic agents. No single agent or combination of agents has resulted in significant prolongation of survival. Therefore, phase II evaluations of new drugs in colorectal carcinoma is of paramount importance. *N*-(Phosphonacetyl)-L-aspartic acid (PALA) is a pyrimidine antimetabolite which has been shown to inhibit aspartate transcarbamylase (1–3). PALA has been shown to have antitumor activity against B16 melanoma, Lewis lung carcinoma, and a variety of transplanted rodent mammary carcinomas (4). It has not been effective against L1210 and P388 mouse leukemias, but it does have activity in xenographs of human adenocarcinomas of the colon transplanted into athymic nude mice (5). Sensitivity of neoplastic cells has been related to the level of aspartate transcarbamylase activity with acquired resistance in tissue culture cells resulting from an increase in enzyme levels (6).

The toxicity of PALA has been studied in dogs, monkeys, and mice, with injury to the gastrointestinal tracts being the primary dose-limiting factors.[3] Gastrointestinal toxicity was especially marked in dogs and monkeys but hepatic toxicity was predominantly noted in mice. This gastrointestinal toxicity has taken the form of mucosal atrophy and hemorrhage with secondary suppurative inflammation when doses of 9.6 g/m² as a single injection or 1.2 g/m²/day × 5 were given to dogs. In monkeys, gastrointestinal epithelial denudation was seen after doses of ≥ 1200 mg/m²/day × 5 days were administered. Neutropenia was noted in monkeys. Both dogs and monkeys manifested other toxic effects of PALA including hyperglycemia, elevation of bilirubin and hepatic enzymes, and mild elevation of creatinine and BUN. CNS toxicity, manifested by the acute onset of sedation, coma, tetany, seizures, respiratory arrests, and in some cases sudden death, has been noted in both the single-dose and the daily-times-five schedules. Doses of 1.2 g/m²/day × 5 days or 2.4 g/m² in a single dose in dogs and doses of 4.8 g/m²/day × 5 days in monkeys caused this acute CNS toxicity.

The study of radiolabeled PALA in animal species demonstrates that the drug is excreted rapidly as the parent compound by the kidneys.[4]

Toxic effects that have been seen in phase I evaluations at the Baltimore Cancer Research Program (7) include occasional nausea and vomiting, skin rash, and diarrhea (at times bloody). The skin toxicity seen was considered the site of major toxicity observed in our study. At a dose of 2.0 g/m² × 5 days, five of five patients treated

[1]Received Aug 2, 1979; revised Sept 28, 1979; accepted Oct 25, 1979.

[2]Baltimore Cancer Research Program, Clinical Oncology Branch, Division of Cancer Treatment, National Cancer Institute, Baltimore, MD.

**Reprint requests to:* David A. Van Echo, MD, Clinical Oncology Branch, Baltimore Cancer Research Program, Division of Cancer Treatment, National Cancer Institute, University of Maryland Hospital, 22 S Greene St, Baltimore, MD 21201.

[3]Clinical brochure: PALA (NSC-224131). Prepared by the Division of Cancer Treatment, National Cancer Institute, Bethesda, MD, Dec 1977, p 24.

[4]Data compiled by the Arthur D. Little, Inc. *In* Clinical brochure cited in footnote 3, p 22.

developed a generalized erythematous, painful, pruritic rash concentrated in the groin and areas of prior cutaneous damage (surgical scars, elbows, etc). The rash was associated with mouth ulcers in two of these patients. Two patients required hospitalization and systemic steroids. Other phase I investigators employing schedules other than the daily-times-five schedule have noted that the major toxic effects of PALA in order of ascending severity are skin rash, mouth ulcers, and bloody diarrhea (8,9). The bloody diarrhea is associated with colonic lesions morphologically indistinguishable from ulcerative colitis by sigmoidoscopy and biopsy (10). In view of this, a dose of 1.5 g/m²/day × 5 days iv should be used for phase II evaluations.

Because of the activity of PALA in animal solid tumors, especially those with slow growth rates, it is of interest to study the effect of PALA in human adenocarcinoma of the colon or rectum.

## MATERIALS AND METHODS

### Patient Population

Twenty-one patients with histologically proven, metastatic recurrent adenocarcinoma of the colon or rectum were treated (*see* table 1).[5] All patients had measurable disease. The median age of the patients was 59.5 years. Thirteen males and eight females were treated. Eighteen of the patients were white. Nine of the patients had liver involvement alone, and two each had local and liver involvement, local and lung involvement, local and lung and liver involvement, or lung and liver involvement. Two patients had single sites of involvement (lung or local), one had liver plus bone involvement, and one had liver plus bone plus lung involvement.

Six patients were previously untreated except for primary surgical intervention at the time of initial presentation. Fifteen patients had received prior chemotherapy. In ten of these patients the therapy was myelosuppressive in nature and included nitrosourea and 5-fluorouracil (5-FU) combinations or 5-FU alone, and in five patients the previous chemotherapy was marrow-sparing (anguidine, 4′-[9-acridinylamino]methanesulfon-*m*-anisidide, or thymidine). Three patients had received irradiation to sites of local involvement in addition to chemotherapy prior to entering this study.

Patients received therapy with PALA at a dose of 1.5 g/m²/day × 5 days iv. The drug was dissolved in 50 ml of dextrose and infused over 15 minutes. Courses of chemotherapy were repeated every 3 weeks. All patients were to receive at least two courses of chemotherapy and, if at the end of these two courses clear progression of the underlying disease had occurred, PALA was discontinued. Patients who were considered to have stabilized while taking PALA continued to receive the drug until clearcut disease progression occurred. Patients were seen at least weekly to assess any problems that may have arisen secondary to their disease or therapy.

Two patients received less than one full course of therapy and are considered unevaluable for response: one died on Day 8 of therapy of complications secondary to tumor progression and one refused to return for the fifth dose of PALA and was lost to followup. The distribution of PALA courses was: one full course (six patients); two courses (four); three courses (six); four courses (two); five courses (one); six courses (one); and eight courses (one).

Patients who developed severe skin or mucosal toxicity were treated with systemic steroids, usually prednisone at a dose of 40–60 mg/day × 7. Upon re-treatment with PALA the steroids were given on the days PALA was administered to prevent or blunt the anticipated toxicity, but the dose of PALA was not decreased unless previous toxicity had been unusually severe.

TABLE 1.—Patient characteristics

| Characteristic | |
|---|---|
| No. of patients entered | 21 |
| No. of patients evaluable | 19 |
| Age (yrs) | |
| Median: 59.5 | |
| Range: 27–75 | |
| Sex | |
| Male | 13 |
| Female | 8 |
| Prior therapy | |
| Chemotherapy | 15 |
| Chemotherapy + radiotherapy | 3 |
| None | 6 |
| Sites of involvement | |
| Liver only | 9 |
| Local + liver | 2 |
| Local + lung | 2 |
| Local + lung + liver | 2 |
| Liver + lung | 2 |
| Lung only | 2 |
| Local only | 2 |
| Liver + bone | 1 |
| Liver + bone + lung | 1 |

## CRITERIA FOR RESPONSE

The criteria used for evaluation of response in this study was that proposed by Moertel (11).

[5]Included are 3 patients previously reported as part of our phase I study of PALA who qualified for this study because of measurable disease and adequate followup, and who received PALA at a dose of at least 1.5 g/m²/day × 5/course.

## RESULTS

### Patient Response

No patient showed regression of measurable disease after treatment with PALA. Eleven patients, however, were able to achieve stabilization of previously progressive disease for periods of 6–31+ weeks. Two patients are considered unevaluable for response because they received less than one full course of therapy.

### Toxicity

Cutaneous toxicity of PALA administration was marked, causing the discontinuation of therapy because of severe rash and/or mucosal toxicity in two patients after one course. Skin toxicity was evaluated according to the following criteria: 0 = no skin abnormality; 1+ = rash involving 10% of the body surface area; 2+ = rash involving 10%–25% of the body surface area; 3+ = rash involving 25%–30% of the body surface area; and 4+ = rash involving > 50% of the body surface area or erythema multiforme. One patient was lost to followup after Day 4 of the first course and is not evaluable for toxicity. Eighteen patients (90%) developed a skin rash related to PALA during the study: two (10%) had a 1+ rash, five (25%) had a 2+ rash, nine (45%) had a 3+ rash, and two (10%) had a 4+ rash. The rash usually began on Day 8 of therapy and lasted for 10 days. The rash did not always occur with the first course of PALA. In five patients it did not occur until the second course, however, all patients who experienced a rash did so by the second course. The rash was initially erythematous and then progressed to scaling and ulceration. It began in intertriginous areas and the face and then proceeded to involve the trunk and arms and legs. From our previous experience (*see* footnote 4) it was known that corticosteroids ameliorated the PALA rash and all patients in this study who developed a ≥ 2+ rash received systemic steroids, usually prednisone at a dose of 40–60 mg/day × 7–10 days. With subsequent PALA therapy it was recommended that corticosteroids be given on the days that PALA was administered and this significantly decreased the cutaneous toxicity in most cases.

Mucositis occurred in seven patients (35%) and was moderately severe in five (25%). It invariably occurred coincident with a skin rash.

Ten patients (50%) developed diarrhea and in five of these patients (25%) it was grossly bloody or strongly guaiac-positive. None of these patients had diverticulitis, a history of colitis or ileitis, or were felt to have tumor involving the lumen of their gastrointestinal tract. Again, diarrhea developed coincident with a skin rash and paralleled it in severity and was ameliorated or prevented by systemic steroids in most cases.

No patient had any myelosuppression during the study: ie, wbc count of 4000/mm$^3$ or platelet count of < 100,000/mm$^3$.

### Survival

The median survival from the start of therapy for 20 patients treated in this group is 14 weeks (range, 1–44+). Ten of the 20 patients have died. No patient suffered a drug-related death.

## DISCUSSION

The dose and schedule used in this study produced significant toxicity and should be considered too high for further phase II studies. A dose of PALA of 1.0–1.25 g/m$^2$/day × 5 days is probably a more reasonable dosage for phase II studies. Major toxicity is cutaneous and mucosal. Gastrointestinal toxicity was manifested by diarrhea which at times was bloody. The drug was discontinued in two patients because of the severity of skin rash and diarrhea. There was no evidence of nausea or vomiting, CNS toxicity, or myelosuppression in this study.

As used in this study, PALA proved ineffective in producing disease regression for patients with metastatic adenocarcinoma of the colon or rectum although several patients achieved stabilization for ≤ 31+ weeks. These results are similar to other preliminary phase II results of the drug in patients with colorectal carcinomas (9). PALA should therefore be added to the list of adequately tested agents that are inactive in patients with colorectal carcinoma.

## REFERENCES


1. Collins KD, and Stark GR. Aspartate transcarbamylase. Interaction with the transition state analogue *N*-(phosphonoacetyl)-L-aspartate. J Biol Chem 246:6599–6605, 1971.
2. Holaday BJ, and Fristrom JW. Phosphonoacetyl-L-aspartate: an aspartate transcarbamylase inhibitor causing larval death and rudimentary wing phenocopies in *Drosphila* melanogaster. Experientia 33:428–430, 1977.
3. Swyryd EA, Seaver SS, and Stark GR. *N*-(Phosphonoacetyl)-L-aspartate, a potent transition state analog inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. J Biol Chem 249:6945–6950, 1974.
4. Johnson RK, Inouye T, Goldin A, et al. Antitumor activity of *N*-(phosphonoacetyl)-L-aspartic acid, a transition-state inhibitor of aspartate transcarbamylase. Cancer Res 36:2720–2725, 1976.
5. Schabel FM, Jr, Laster WR, Jr, and Rose WC. Experimental Chemotherapy and Tumor Cell Kinetics: Overview of Experimental Tumor Systems in Treatment of Lung Cancer (Muggia FM, and Rozencweig M, eds). New York, NY, Raven Press, 1977, vol 11, pp 15–35.
6. Johnson RK, Swyryd EA, and Stark GR. Effects of *N*-(phosphonoacetyl)-L-aspartate on murine tumors and normal

tissues in vivo and in vitro and the relationship of sensitivity to rate of proliferation and level of aspartate transcarbamylase. Cancer Res 38:371–378, 1978.
7. ERLICHMAN C, STRONG J, WIERNIK P, ET AL. Phase I trial of PALA (*N*-phosphonoacetyl)-L-aspartate). Proc Am Assoc Cancer Res and ASCO 20:314, 1979.
8. VALDIVIESO M, MOORE EC, LOO TL, ET AL. Phase I clinical study of *N*-(phosphonoacetyl)-L-aspartate (PALA, NSC 224131). Proc Am Assoc Cancer Res and ASCO 20:187, 1979.
9. OHNUMA T, HART R, ROBOZ J, ET AL. Clinical and pharmacological studies with phosphonoacetyl-L-aspartate (PALA). Proc Am Assoc Cancer Res and ASCO 20:344, 1979.
10. GRALLA RJ, CASPER ES, GOLBEY RB, ET AL. Phase I and preliminary phase II studies with *N*-(phosphonacetyl)-L-aspartic acid (PALA). Proc Am Assoc Cancer Res and ASCO 20:115, 1979.
11. MOERTEL CG. Chemotherapy of gastrointestinal cancer. Clin Gastroenterol 5:777, 1976.

# Phase II Study of Tamoxifen in Patients With Advanced Renal Cell Carcinoma [1,2,3]

Brief Reports and Preliminary Communications

John H. Glick,* Alan Wein, Susan Torri, Jane Alavi, David Harris, and Harvey Brodovsky [4,5]

Hormonal therapy for advanced renal cell carcinoma using progestational agents is a well-established approach to the management of metastatic disease, although objective responses are not common (1–3). The nonsteroidal antiestrogen, nafoxidine, produced objective remissions in an estrogen-induced hamster carcinoma (4); this led the Eastern Cooperative Oncology Group (ECOG)[6] to conduct a Phase II trial of nafoxidine in metastatic renal cell cancer which produced a response rate of 15%. However, Feun and Oishi (5) reported only a 5% objective response rate in 21 patients with inoperable or metastatic adenocarcinoma of the kidney treated with nafoxidine.

Recently, the antiestrogenic agent, tamoxifen, has been introduced for the palliative treatment of advanced breast carcinoma. Tamoxifen is a nonsteroidal derivative of triphenylethylene which specifically binds to estrogen-receptor sites in target tissue. The demonstration of estrogen receptors in human renal cell carcinoma (6,7) raised the possibility that this tumor may be a target tissue for estrogen and therefore sensitive to antiestrogen therapy. This observation, and recent reports of occasional nafoxidine-induced regression (5), led to this phase II trial of tamoxifen in metastatic renal cell carcinoma.

[1]Received Aug 8, 1979; revised Oct 22, 1979; accepted Oct 29, 1979.

[2]Supported in part by Public Health Service grants CA-16420, CA-45055, and CA-06927 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare, and by ICI Americas, Inc, Wilmington, DE.

[3]Tamoxifen was supplied through the courtesy of ICI Americas, Inc, Wilmington, DE.

[4]Hematology-Oncology Section, Department of Medicine (*J. H. Glick, S. Torri,* and *J. Alavi*), and Division of Urology (*A. Wein*), Hospital of the University of Pennsylvania, American Oncologic Hospital (*D. Harris*), and Thomas Jefferson University Hospital (*H. Brodovsky*), Philadelphia.

[5]We thank Dr. Robert Catalano and Catherine Duggery for their assistance in data collection, and Deborah Altieri and Patricia Lafferty for secretarial assistance.

[6]ECOG protocol No. 0269. Oral Provera, hydroxyurea and U-11, 100A in renal cell carcinoma. Group minutes, Sept 1973.

**Reprint requests to:* John H. Glick, MD, Hematology-Oncology Section, Department of Medicine, Hospital of the University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104.

## MATERIALS AND METHODS

Patients selected for this study had histologically confirmed renal cell carcinoma not amenable to surgical cure (primary inoperable, recurrent, or metastatic disease). Patients who had received prior hormonal therapy (progestational agents) or prior chemotherapy were eligible for the study provided disease progression was documented and at least 2 weeks had elapsed since the cessation of hormonal therapy or chemotherapy. Only patients with measurable disease were eligible. Written informed consent was obtained.

Tamoxifen was administered at a dose of 10 mg orally twice a day for 8 weeks. Patients were examined at 2–4-week intervals for objective and subjective evaluation of tumor response. If the patient had achieved a partial response or disease stabilization at 8 weeks, tamoxifen was continued until progressive disease was documented and the drug was then stopped. If the patient had progressive disease at 8 weeks, tamoxifen was discontinued. Therapeutic responses were evaluated in the following manner: complete response was defined as the disappearance of all clinical evidence of cancer; partial response was defined as at least a 50% surface area reduction of measurable lesions for a minimum of 4 weeks; disease stabilization (no change) was defined as a $< 50\%$ reduction in measurable tumor without an increase in size of any known malignant disease or the appearance of new lesions; and progressive disease was defined as the appearance of any new lesion or a 25% increase in measurable disease.

The ECOG performance status was used to evaluate subjective response.

## RESULTS

Fifteen patients with metastatic renal cell carcinoma have been entered in this study as of June 1, 1979. Table 1 describes some patient characteristics of the 15 patients, all of whom are evaluable. None of the patients had determinations of estrogen- or progesterone-binding proteins on either their primary or metastatic lesions. No complete or partial responses were seen in the 15 patients with metastatic renal cell carcinoma who were treated with tamoxifen. After 8 weeks of tamoxifen, therapy was

TABLE 1.—Patient characteristics

| Characteristic | No. of patients |
|---|---|
| Total | 15 |
| Sex | |
| Males | 11 |
| Females | 4 |
| Median age in yrs: 64 | |
| Prior therapy | |
| Nephrectomy | 10 |
| Progestins | 5 |
| Chemotherapy | 2 |
| ECOG performance status | |
| 0–1 (ambulatory) | 11 |
| 2–3 (bedridden) | 4 |
| Sites of metastatic disease | |
| Lung | 6 |
| Lung, liver | 2 |
| Lung, liver, bone | 2 |
| Liver, soft tissue | 1 |
| Lung, bone | 1 |
| Bone | 2 |
| Soft tissue (nodes) | 1 |

discontinued in 12 patients because of progressive disease. Of the 12 patients with progressive disease, none experienced any subjective improvement and seven deteriorated symptomatically with worsening performance status during the 8-week trial with tamoxifen. During the initial 8-week treatment period, three patients showed no change (disease stabilization) in their measurable lesions and therefore were continued in the study until they developed progressive disease. None of the three patients who achieved disease stabilization had received prior hormonal therapy or chemotherapy, but all had undergone prior nephrectomy. The duration of disease stabilization was 364, 286, and 126 days in the three patients, all of whom have subsequently developed progressive disease. All three of these patients are alive with a median survival time of 358+ days, whereas nine of the 12 patients with progressive disease have died (median survival time, 155 days). Side effects were extremely mild and caused no patient to be withdrawn from the study. Specifically, hypercalcemia and tumor flares were not observed.

## DISCUSSION

Tamoxifen does not appear to be a useful agent in the treatment of advanced metastatic renal cell carcinoma. Our series did not demonstrate a single objective response, and only three of 15 patients (20%) achieved disease stabilization. It is recognized that disease stabilization may occur spontaneously and last for long periods of time in the natural history of renal cell carcinoma (8). Therefore, it is difficult to attribute significant benefit to tamoxifen in our patients who achieved disease stabilization, although their survival appears to be prolonged.

These data confirm prior reports of the relative ineffectiveness of antiestrogens in the treatment of renal cell cancer. Nafoxidine produced response rates of 5% and 15% in two reports. Recently, the Southwest Oncology Group reported the preliminary results from a phase II trial of tamoxifen in renal cell cancer. Of 49 patients, they noted two patients with partial remissions and 14 patients (28%) with disease stabilization (9). However, the median survival time was only 15.5 weeks for those patients achieving partial remission or disease stabilization in this series.

Although estrogen and progesterone receptors have been reported in renal cell carcinoma (10), none of our patients had measurements of receptors on either their primary or metastatic lesions. However, the usefulness of receptor measurements in renal cell carcinoma remains doubtful in view of the extremely small fraction of patients reported to respond to progestin therapy despite the presence of progesterone receptors in 90% of the specimens assayed.

In our experience, the antiestrogenic agent, tamoxifen, has proven to be of little therapeutic value in the management of advanced renal cell carcinoma.

## REFERENCES


1. LOKICH JJ, and HARRISON JH. Renal cell carcinoma: natural history and chemotherapeutic experience. J Urol 144:371–374, 1975.
2. ALBERTO P, and SENN HJ. Hormonal therapy of renal carcinoma alone and in association with cytostatic drugs. Cancer 33:1226–1229, 1974.
3. MORALES A, KIRULUTA G, and LOTT S. Hormones in the treatment of metastatic renal cancer. J Urol 114:692–693, 1975.
4. BLOOM HJG, ROE FJC, and MITCHLEY BCB. Sex hormones and renal neoplasia. Inhibition of tumor of hamster kidney by an estrogen antagonist, an agent of possible therapeutic value in man. Cancer 20:2118–2124, 1967.
5. FEUN LG, and OISHI N. Letter: Phase II study of nafoxidine in the therapy for advanced renal carcinoma. Cancer Treat Rep 63:149–150, 1979.
6. GIULIANI L, PESCATORE D, GIBERTI C, ET AL. Usefulness and limitation of estrogen receptor protein (ERP) assay in human renal cell carcinomas. Eur Urol 4:342–347, 1978.
7. BOJAR H, WITTLIFF J, and BALZER K. Properties of specific estrogen-binding components in human kidney and renal carcinoma. Acta Endocrinol 193:51, 1975.
8. HOLLAND JM. Natural history and staging of renal cell cancer. CA 25:121–133, 1975.
9. AL-SARRAF M. The clinical trial of tamoxifen in patients with advanced renal cell cancer: a SWOG study. Proc Am Assoc Cancer Res and ASCO 20:378, 1979.
10. FRIEDMAN MA, HOFFMAN PG, and JONES HW. The clinical value of hormone receptor assays in malignant disease. Cancer Treat Rev 5:185–194, 1978.

# Phase II Evaluation of 4′-(9-Acridinylamino)-methanesulfon-*m*-anisidide (AMSA) in Patients With Non-Small Cell Lung Cancer [1,2]

Brief Reports and Preliminary Communications

**Ephraim S. Casper,* Richard J. Gralla, David P. Kelsen, Ronald B. Natale, Peter Sordillo, and Alan Houghton** [3,4]

4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA) is an acridine derivative developed by Cain and Atwell (1). It possesses antitumor activity in L1210 and P388 murine leukemias, and in several solid tumor models including B16 melanoma, C3H and $CD8F_1$ mammary tumors, and colon tumors 26 and 38.[5] AMSA binds to DNA, removing and reversing the supercoiling of closed circular duplex PM2 DNA, suggesting that it is an intercalating agent (2).

Phase I studies indicate that at a dose of 90–120 mg/m² every 21–28 days, AMSA produces consistent, but acceptable, myelosuppression with mild nonhematologic toxicity (3,4). The drug is metabolized principally within the liver and excreted in the bile; however, urinary excretion increases in the presence of hepatic dysfunction (5). Antitumor activity has been noted in acute leukemia (6,7), melanoma (7,8), Hodgkin's disease (9), chronic lymphocytic leukemia (9), liposarcoma (4), squamous carcinomas of the esophagus (9) and unknown primary site (4), and adenocarcinomas of the ovary (3), breast (6), unknown primary site (4), and lung (7). Because of the reported activity of AMSA in this variety of malignancies including squamous carcinomas and adenocarcinomas of the lung, we have conducted a phase II trial of this new agent in patients with non-small cell lung cancer.

## MATERIALS AND METHODS

Twenty-nine patients with a histologically or cytologically confirmed diagnosis of non-small cell carcinoma of the lung were entered in the study. Criteria for eligibility included the presence of measurable lesions which had not been previously irradiated, a performance status (Karnofsky scale) of $\geq 50\%$, normal hepatic and renal functions, and a wbc count of $> 4000/mm^3$ and platelet count of $> 100{,}000/mm^3$.

Twenty-two patients were evaluable for therapeutic response, including 18 with adenocarcinoma, three with epidermoid carcinoma, and one with large cell carcinoma. The median age was 61 years (range, 38–73) and the median performance status was 70% (range, 50%–90%). The male to female ratio was 14 to eight. Seven patients had received no prior chemotherapy; four of these seven patients had previously been treated with radiotherapy. Twelve patients had been previously treated with chemotherapy alone and three others had been previously treated with both radiotherapy and chemotherapy. Seven of the original 29 patients did not receive adequate trials of AMSA. Chemotherapy was interrupted in five patients so that a course of radiotherapy could be given to control local complications of disease; two patients died of progressive disease before the development of myelotoxicity from AMSA.

AMSA was provided by the Division of Cancer Treatment, National Cancer Institute (Bethesda, MD) as a 75-mg solution in 1.5 ml of *N*,*N*-dimethylacetamide. This solution was then diluted in 0.0353 M L-lactic acid to a total of 15 ml, yielding a concentration of 5 mg/ml. The appropriate dose was further diluted in 5% dextrose in water to a volume of 250 ml and was administered as an iv infusion over 45 minutes through a running iv line. The drug was given every 21 days at a dose of 90 mg/m² for patients with a performance status of 50%–70% and at a dose of 120 mg/m² for patients with a performance status of $\geq 80\%$. If the wbc count fell to $< 3000/mm^3$, the dose was decreased by 30 mg/m²; if there was no myelosup-

[1]Received Aug 8, 1979; revised Oct 30, 1979; accepted Nov 7, 1979.

[2]Supported in part by contract N01-CM-97274 (Division of Cancer Treatment) and Public Health Service grant CA-05826 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare. Dr. Kelsen is a Landrum/Karnofsky Scholar and Dr. Gralla is an American Cancer Society Junior Faculty Clinical Fellow.

[3]Developmental Chemotherapy (*E. S. Casper, R. J. Gralla, P. Sordillo,* and *A. Houghton*) and Solid Tumor (*D. P. Kelsen* and *R. B. Natale*) Services, Department of Medicine, Memorial Sloan-Kettering Cancer Center, Cornell University Medical College, New York, NY.

[4]We wish to thank Polly Hodge for preparation of the manuscript and Jane Howard and Marc Serber for assistance with analysis of the data.

[5]Screening data from the Drug Evaluation Branch, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD.

**Reprint requests to:* Ephraim S. Casper, MD, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10021.

pression, the dose was escalated by 30 mg/m². An adequate trial was defined as treatment producing myelotoxicity (wbc count < 3000/mm³ or platelet count < 50,000/mm³) in the presence of unequivocal evidence of progressive disease, or treatment with at least two doses of ≥ 120 mg/m² of AMSA.

Criteria for therapeutic responses were as follows: *complete response*—the disappearance of all evidence of tumor for at least 1 month; *partial response*—a ≥ 50% reduction in the sum of the products of the two greatest perpendicular dimensions of all measurable lesions for at least 1 month; *minor regression*—a ≥ 25% reduction of all measurable lesions, but less than that required for a partial response; and *stable disease*—a < 25% reduction of all measurable lesions in the absence of unequivocal progression of disease.

## RESULTS

*Therapeutic responses.*—No major objective responses occurred. One patient with adenocarcinoma who had failed previous chemotherapy had a minor regression with reduction in a metastatic mandibular mass and marked diminution of pain in that area. She died unexpectedly of an unknown cause during her third month of treatment. Six patients with adenocarcinoma, two of whom had received no prior chemotherapy, had stable disease lasting 2, 3, 3, 5, 7+, and 8+ months.

*Toxicity.*—The median number of doses administered to each patient was three (range, one to 11). Myelosuppression, primarily leukopenia, was the major toxic effect observed. The median wbc count nadir was 2800/mm³ (range, 1200–9000) and the median platelet count nadir was 209,000/mm³ (range, 107,000–465,000). The median wbc and platelet count nadirs occurred on Day 15 of treatment for outpatients who were followed weekly and on Days 7–10 of treatment for inpatients who were followed more closely. It is interesting that the five patients who did not develop a wbc count nadir of < 3500/mm³ included one patient with the minor regression and three patients with stable disease. The doses were not escalated in four of these patients because of early death, severe mucositis, severe nausea and vomiting, and general clinical deterioration, respectively. The other patient without leukopenia had rapid progression of disease in spite of two doses of AMSA at 120 and 150 mg/m². During their courses of treatment, 19 patients had a decrease in hemoglobin concentration ranging from 1.1 to 3.6 g/dl (median, 2.5 g/dl). This fall in hemoglobin tended to be cumulative, although the hemoglobin concentration fell to < 8 g/dl in only four patients; it was most profound in patients with rapid progression of disease. Six patients developed mild nausea and one patient had severe nausea and vomiting on the day of drug administration. Three patients had mild to moderate mucositis. Neither phlebitis nor hepatic or renal dysfunction was observed. Approximately 80% of the patients complained of a vague feeling of malaise on the day of drug administration.

## DISCUSSION

It has been reported that AMSA has antitumor activity in several human malignancies including two minor regressions in four patients with adenocarcinoma of the lung (7). The present study was conducted in patients having a good performance status. One third of these patients had received no prior chemotherapy. Toxicity, primarily leukopenia, did occur but no major objective responses were seen among the 22 evaluable patients. Eighteen of the patients had adenocarcinoma. At 95% confidence limits, the predicted true major response rate is < 16% for patients with adenocarcinoma of the lung. We have not seen a significant difference in response rate among the subclassifications of non-small cell bronchogenic carcinoma in previous studies at this institution (10,11). Including all 22 patients, the predicted response rate of AMSA is < 14% ($P = 0.05$) for patients with non-small cell lung cancer.

We conclude that AMSA, in the dose and schedule given to this population of patients, is not active as a single agent in the treatment of non-small cell carcinoma of the lung.

## REFERENCES


1. Cain BF, and Atwell GJ. The experimental antitumour properties of three congeners of the acridylmethanesulphonanilide (AMSA) series. Eur J Cancer 10:539–549, 1974.
2. Waring MJ. DNA-binding characteristics of acridylmethanesulphonanilide drugs: comparision with antitumour properties. Eur J Cancer 12:995–1001, 1976.
3. Von Hoff DD, Howser D, Gormley P, et al. Phase I study of methanesulfonamide, *N*-[4-(9-acridinylamino)-3-methoxyphenyl]-(*m*-AMSA) using a single-dose schedule. Cancer Treat Rep 62:1421–1426, 1978.
4. Van Echo DA, Gormley P, Lichtenfeld JL, et al. A phase I study of *m*-AMSA. Proc Am Assoc Cancer Res and ASCO 20:422, 1979.
5. Hall SW, Benjamin RS, Legha SS, et al. Clinical pharmacokinetics of the new antitumor agent AMSA. Proc Am Assoc Cancer Res and ASCO 20:175, 1979.
6. Legha SS, Bodey GP, Keating MJ, et al. Early clinical evaluation of acridinyl-amino-methanesulfon-*m*-anisidide (AMSA) in patients with advanced breast cancer and acute leukemia. Proc Am Assoc Cancer Res and ASCO 20-416, 1979.
7. Legha SS, Gutterman JU, Hall SW, et al. Phase 1 clinical investigation of 4′-(9-acridinylamino) methanesulfon-*m*-anisidide (NSC 249992), a new acridine derivative. Cancer Res 38:3712–3716, 1978.
8. Hall SW, Benjamin RS, Legha SS, et al. AMSA: a new acridine derivative with activity against metastatic melanoma. Proc Am Assoc Cancer Res and ASCO 20:372, 1979.
9. Goldsmith MA, Bhardwaj S, Ohnuma T, et al. Phase I study of *m*-AMSA in patients with solid tumors and leukemias. Proc Am Assoc Cancer Res and ASCO 20:344, 1979.

10. GRALLA RJ, RAPHAEL BG, GOLBEY RB, ET AL. Phase II evaluation of vindesine in patients with non-small cell carcinoma of the lung. Cancer Treat Rep 63:1343–1346, 1979.

11. CASPER ES, GRALLA RJ, and GOLBEY RB. Vindesine (DVA) and *cis*-dichlorodiamminoplatinum II (DDP) combination chemotherapy in non-small cell lung cancer. Proc Am Assoc Cancer Res and ASCO 20:337,1979.

# Phase II Evaluation of *N*-(Phosphonacetyl-L-aspartic Acid (PALA) in Patients With Advanced Colorectal Carcinoma [1,2,3]

**Brief Reports and Preliminary Communications**

**David S. Carroll, Richard J. Gralla, and Nancy E. Kemeny[4,5,*]**

*N*-(Phosphonacetyl)-L-aspartic acid (PALA) is an inhibitor of aspartate transcarbamylase which catalyzes the reaction between carbamyl phosphate and L-aspartic acid, the first step in de novo pyrimidine nucleotide biosynthesis (1). Cell culture studies with normal primate fibroblasts and human and rodent colonic carcinoma cells revealed that normal cells respond to PALA by growth arrest, presumably by shifting into a $G_1$ state, while the neoplastic cells demonstrated continued cell division with limited survival (2). In transplantable rodent tumors, PALA is more effective against certain solid tumors, specifically Lewis lung carcinoma and B16 melanoma, than against L1210 and P388 leukemias (3). There is evidence that rapidly growing murine leukemias have higher levels of aspartate transcarbamylase than the more slowly growing solid tumors which are sensitive to PALA (4). Several phase I trials demonstrated gastrointestinal and cutaneous toxicity without significant myelosuppression (5–9). Between November 1978 and May 1979, a phase II trial of PALA was conducted in patients with metastatic colorectal carcinoma at Memorial Sloan-Kettering Cancer Center (MSKCC).

## MATERIALS AND METHODS

All patients with measurable and histologically confirmed advanced colorectal carcinoma were eligible for the study if their Karnofsky performance status (PS) was $> 50$ and if their serum creatinine level was $< 2$ mg/100 ml. The initial dose of PALA was 4.5 g/m² for patients with a PS of $\geq 70$ and 3.7 g/m² for a PS of $< 70$. The dose of PALA was diluted in 50 mg of normal saline and administered iv over 30–60 minutes weekly. Diarrhea and rash were graded according to the criteria presented in table 1. In the presence of grade 2+ toxicity (table 1), the drug was withheld until symptoms subsided, and then treatment was resumed with a 20% reduction in dose. When no toxicity developed after two weekly doses, the subsequent dose was increased by 20%.

An adequate trial of PALA was defined as the administration of at least three doses without toxicity or two doses with grade 3+ cutaneous and/or gastrointestinal toxicity (table 1). A partial remission denoted a $> 50\%$ decrease in the sum of the products of the two largest perpendicular diameters of all measurable lesions. A minor response was a $> 25\%$ decrease in all measurable disease, and stabilization was a $< 25\%$ change in measurable disease in the absence of new lesions.

The median age for the 21 patients entered in this study was 60 years (range, 42–71), and the median performance status was 80 (range, 50–100). There were 11 males and ten females. All but one patient had received prior chemotherapy; 71% had had two previous drug trials. Eighty-four percent of patients had received methyl-CCNU and 5-fluorouracil (5-FU), 38% had received vincristine and streptozotocin, and 19% had received vindesine. Three patients (14%) had received prior radiotherapy. The median time from diagnosis of metastatic disease to entry in this study was 17 months. Sites of metastatic disease were: liver (16 patients; 76%), lung (11; 52%), pelvic mass (three; 14%), lymph nodes (two; 10%), vulva (one; 5%), and vagina (one; 5%).

## RESULTS

Nineteen of the 21 patients who were entered in the study received an adequate trial of PALA. One of the two inadequate patients died from progressive disease after a single dose, and the other demonstrated only grade 1+ dermatologic toxicity after two doses. For the evaluable patients the median total dose of PALA was 18.8 g/m² (range, 9.0–89.3 g/m²) and the median number of doses was five (range, two to 20). Eighteen patients

[1]Received Aug 9, 1979; revised Oct 31, 1979; accepted Nov 13, 1979.

[2]Supported in part by Public Health Service grant CA-05826 and contract N01-CM-97274 (Division of Cancer Treatment) from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare. Dr. Carroll is a fellow in Medical Oncology at Memorial Sloan-Kettering Cancer Center (New York, NY).

[3]PALA was supplied by the Division of Cancer Treatment, National Cancer Institute (Bethesda, MD).

[4]Solid Tumor (*D. S. Carroll* and *N. E. Kemeny*) and Developmental Chemotherapy (*R. J. Gralla*) Services, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY.

[5]We wish to acknowledge the technical assistance of Jane Howard and the secretarial assistance of Cathie Cassidy.

**Reprint requests to:* Nancy E. Kemeny, MD, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10021.

TABLE 1.—Gastrointestinal and cutaneous toxicity in 19 patients treated with PALA

| Grade | Gastrointestinal toxicity | No. of patients (%) | Cutaneous toxicity | No. of patients (%) |
|---|---|---|---|---|
| 0 | No gastrointestinal symptoms | 6 (31) | No skin abnormality | 3 (16) |
| 1+ | Mild diarrhea, not interfering with normal activities | 5 (26) | Rash involving <10% of body surface area | 5 (26) |
| 2+ | Uncomfortable moderate diarrhea interfering with normal activities | 3 (16) | Rash involving 10%–25% of body surface area | 6 (31) |
| 3+ | Profuse, painful, bloody diarrhea | 4 (21) | Rash involving 25%–50% of body surface area | 4 (21) |
| 4+ | Diarrhea causing fluid and electrolyte imbalance requiring hospitalization | 1 (5) | Rash involving >50% of body surface area, or erythema multiforme | 1 (5) |

received 4.5 g/m$^2$ initially, of whom five had dose escalations and four had a decrease in dose. Three patients initially received the low dose, 3.7 g/m$^2$, and only one had an escalation of the dose.

Two patients with pulmonary metastases achieved minor responses for 3.5 and 7.0+ months. Chest radiographs demonstrated a 30% and 40% decrease in measurable pulmonary disease. Ten patients had stable metastatic disease for a median duration of 6 weeks (range, 3–8) during the course of treatment, and seven developed progressive disease.

Rash and diarrhea were the most frequent adverse effects of therapy. Rash, observed in 85% of cases, was an erythematous macular and often pruritic eruption which could be seen on any part of the trunk, arms, legs, face, or genitalia. It was most frequently seen 3–5 days after the second weekly dose of PALA and subsided in 5–10 days after discontinuing the drug. Areas of hyperpigmentation often remained after other signs of cutaneous toxicity had cleared. In the patient demonstrating the most severe cutaneous toxicity, the rash involved 75% of the body surface area and persisted for 15 days.

Diarrhea occurred in 70% of patients and the time course was similar to that described for the cutaneous toxicity. Mucositis (42%) ranged from minor angular cheilitis to a prolonged diffuse hemorrhagic inflammation of the mouth and pharynx. Six patients (31%) complained of recurrent eye irritation, and mild conjunctival erythema was observed in one half of these patients. The wbc count nadir was >4000 cells/mm$^3$ in 17 patients and 3700 and 2200 cells/mm$^3$ in the remaining two patients. Platelet counts of < 150,000 cells/mm$^3$ were not observed.

Neurologic symptoms were observed in four patients (21%). Marked confusion was present in each case, and two patients had generalized seizures. The symptoms developed after two to four doses of PALA. In one patient confusion began during infusion of the drug, and in the other patients symptoms were observed 1–4 days after the last dose. Metastatic intracranial disease was demonstrated in only one of these patients. In another patient, who also had grade 4+ gastrointestinal and cutaneous toxicity, marked confusion persisted for 3 weeks and gradually improved over the following 10 days. When the potential neurotoxic side effects of PALA became evident, serial electroencephalograms (EEGs) were obtained in nine patients receiving PALA, and a single tracing was obtained on each of the additional six patients. Moderate bilateral slowing and disorganization were observed in all patients who were having neurologic symptoms, but similar EEG abnormalities developed in some patients who received the drug and showed no signs of neurotoxicity. One patient who received 20 doses of PALA had progressive deterioration on four EEGs, but remained completely free from clinical neurotoxicity and had a normal computerized transaxial tomogram. Three patients developed transient complaints of perioral paresthesias, dizziness, and mild confusion during the drug infusion, and these symptoms subsided when the rate of the drug administration was decreased.

Toxicity was not influenced by hepatocellular disease; the two patients with extensive liver disease, ie, alkaline phosphatase > 500 (normal level, 40–80), had only grade 1–2+ toxicity. Four patients had creatinine levels of ≥ 1.5 but < 1.8 mg/100 ml, and all of these patients had grade 3–4+ toxicity, while of the 17 patients with creatinine levels of < 1.5 mg/100 ml, five had grade 3–4+ toxicity ($P$ = 0.026, Fisher's exact test).

## DISCUSSION

Twenty-one patients with metastatic colorectal carcinoma were treated with weekly injections of PALA. Rash and diarrhea were the most common and limiting side effects. The incidence of serious toxicity could not be predicted on the basis of liver function tests but was increased in patients with a creatinine level of ≥ 1.5 mg/100 ml.

The neurologic toxicity observed in 21% of the patients

was seen in one patient with intracerebral structural lesions and in other patients in whom no such lesions could be found. Similar results were seen in other phase II trials with PALA at MSKCC (6). It is possible that the structural lesions present in some of the patients facilitated entry of PALA into the brain with subsequent encephalopathy and lowering of the seizure threshold. Although abnormal EEGs were found in all patients who developed neurotoxicity, no EEG abnormality was predictive of subsequent neurologic symptoms.

Minor responses were seen in two of 19 adequately treated patients. It should be emphasized that the majority of patients in this study were heavily pretreated, although a minor response of 7+ months was seen in a patient who had undergone three prior chemotherapy trials. With a predicted response rate of $< 15\%$ ($P = 0.05$) (10), we conclude that PALA by this dose and schedule is not an effective single agent in previously treated patients with colorectal carcinoma.

## REFERENCES


1. Swyryd EA, Seaver SS, and Stark GR. *N*-(Phosphonacetyl)-L-aspartate, a potent transition state analog inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. J Biol Chem 249:6945–6950, 1974.
2. Tsuboi KK, and Kwong LK. Antiproliferative agents and differential survival between normal and cancer cells. Cancer Res 38:3745–3750, 1978.
3. Johnson RK, Inouye T, Goldin A, et al. Antitumor activity of *N*-(phosphonacetyl)-L-aspartic acid, a transition-state inhibitor of aspartate transcarbamylase. Cancer Res 36:2720–2725, 1976.
4. Johnson RK, Swyryd EA, and Stark GR. Effects of *N*-(phosphonacetyl)-L-aspartate on murine tumors and normal tissues in vivo and in vitro and the relationship of sensitivity to rate of proliferation and level of aspartate transcarbamylase. Cancer Res 38:371–378, 1978.
5. Valdivieso M, Moore EC, Loo TL, et al. Phase I clinical study of *N*-(phosphonacetyl)-L-aspartate (PALA, NSC 224131). Proc Am Assoc Cancer Res and ASCO 20:187, 1979.
6. Gralla RJ, Casper ES, Golbey RB, et al. Phase I and preliminary phase II studies with *N*-(phosphonacetyl)-L-aspartic acid (PALA). Proc Am Assoc Cancer Res and ASCO 20:115, 1979.
7. Erlichman C, Strong J, Wiernik P, et al. Phase I trial of PALA (*N*-phosphonacetyl-L-aspartate). Proc Am Assoc Cancer Res and ASCO 20:314, 1979.
8. Ohnuma T, Hart R, Roboz J, et al. Clinical and pharmacological studies with phosphonacetyl-L-aspartate (PALA). Proc Am Assoc Cancer Res and ASCO 20:344, 1979.
9. Ervin TJ, Blum RH, and Cannelos GP. *N*-Phosphonacetyl-L-aspartate (PALA), phase I trial. Proc Am Assoc Cancer Res and ASCO 20:200, 1979.
10. Gehan EA. The determination of the number of patients required in a preliminary and a follow up trial of a new chemotherapeutic agent. J Chronic Dis 13:346–353, 1961.

# LETTERS

## Cardiomyopathy in Adults After Combination Adriamycin and DTIC

A recent report (1) referred to an unexpected high incidence of cardiomyopathy (CMP) in children with soft tissue and bony sarcomas treated with Adriamycin (ADR) and DTIC (Dacarbazine) plus vincristine or the epipodophyllotoxin derivative, VM-26.

The occurrence of CMP was 40% in ten patients receiving 300–540 mg/m² of ADR with various doses of other drugs. The mean and median total dose of ADR to CMP was 450 mg/m², without correlation with DTIC total doses. The median time to onset of CMP after starting ADR was between 8 and 11 months.

We recently treated 19 adult patients with locally advanced or recurrent soft tissue sarcomas with the combination ADR plus DTIC as adjuvant chemotherapy after surgery with or without local radiation therapy (2).

None of the 19 patients, all of whom had regular followup for at least 1 year after stopping the treatment, developed CMP. The treatment was within the same dose schedule as in Smith's experience (1): 60 mg/m² iv of ADR on Day 1 as a bolus injection, and 250 mg/m²/day iv of DTIC for 5 consecutive days. If the patient was concomitantly treated with radiotherapy the ADR was reduced by 45 mg/m² during two initial courses. No patient was receiving radiotherapy to the mediastinal-pericardial area. Patients did receive ADR at a minimum dose of 135 mg/m² and a maximum dose of 550 mg/m². Ten of the 19 patients had between 400 and 500 mg/m² of ADR; seven of them had $> 500$ mg/m².

Side effects were in the expected range and intensity, similar to what was previously reported after the treatment of patients with advanced disease with the same ADR plus DTIC combination (3).

Our clinical observation is also confirmed by the serial STI recording of cardiac function through the polygraphic examination. In fact, we did record—at the critical level of 400–550 mg/m² ADR total dose—alterations of PEP/LVET ratio (median values: 0.385 for doses between 400 and 500 mg/m² and 0.397 for doses $> 500$ mg/m²), but they are very close to the variations observed after ADR in a previously reported experience (4) and the differences are not significant at the statistical analysis.

In this regard we do not consider the ADR plus DTIC combination to be synergistic in respect to drug-related cardiotoxicity. Our opinion is confirmed by previous experience with the same effective polychemotherapy in the treatment of advanced soft tissue sarcomas (2) and from the recent review concerning ADR cardiotoxicity performed on 1273 retrospectively studied cases by the European group of Praga and co-workers (5,6). In their paper, the authors reported that the incidence of definite ADR cardiomyopathy is 1.8% among 92 patients treated with ADR plus DTIC.

Since there are no available reports of the cardiotoxicity of DTIC in the literature (7), we believe that other co-factors which might explain the reason for the increased cardiotoxicity seen in the Smith study should be investigated.

### REFERENCES


1. Smith PJ, Ekert H, Waters KD, et al. Letter: High incidence of cardiomyopathy in children treated with Adriamycin and DTIC in combination chemotherapy. Cancer Treat Rep 61:1736–1738, 1977.
2. Beretta G. Medical treatment of bone and soft tissue sarcomas. (Abstr 80) *In* Proceedings of the 3rd International Congress of Oncology, Teheran, Iran, Mar 3–8, 1978.
3. Beretta G, Bonadonna G, Bajetta E, et al. Combination chemotherapy with DTIC (NSC-45388) in advanced malignant melanoma, soft tissue sarcomas, and Hodgkin's disease. Cancer Treat Rep 60:205–212, 1976.
4. Villani F, Beretta G, Guindani A, et al. Adriamycin cardiotoxicity in man: effects of pretreatment with beta-methyldigoxin. A polygraphic study. Chemotherapy 8:490–499, 1976.
5. Praga C, Beretta G, Vigo PL, et al. Adriamycin-cardiotoxicity: a survey of 1273 patients. Cancer Treat Rep 63:827–834, 1979.
6. Praga C, Beretta G, and Labianca R. Cardiac toxicity from antitumor therapy. Oncology, 1980. In press.
7. Ghione M. Cardiotoxic effects of antitumor agents. Cancer Chemother Pharmacol 1:25–34, 1978.



Gianni Beretta
Medical Oncology Department
S. Carlo Borromeo Hospital
20153 Milan, Italy

Fabrizio Villani
Cardiology Unit
National Cancer Institute
20133 Milan, Italy


September 4, 1979

## Report on Drug-Induced Pericarditis

Drug-induced pericarditis is rare. It has been reported in association with the administration of actinomycin D (1,2), bleomycin (3), and procainamide (4). Carditis with pericarditis has been reported with daunorubicin therapy and with high-dose cyclophosphamide administration (5). Prior mediastinal irradiation was a contributing factor in some cases (1–3). We have recently observed three epi-

sodes of acute pericarditis in two patients receiving combination chemotherapy with cyclophosphamide, vinblastine, actinomycin D, bleomycin, and *cis*-dichlorodiammineplatinum(II) (cisplatin) for advanced testicular cancer.

## CASE REPORTS

### Patient 1

Patient 1, a 33-year-old white male with teratocarcinoma, choriocarcinoma, and a persistent elevation of serum human chorionic gonadotropin (HCG)-beta following left radical orchiectomy and retroperitoneal lymph node dissection, was admitted to the hospital for chemotherapy. On November 16, 1978, following adequate hydration he received 1.4 g of cyclophosphamide, 2.4 mg of actinomycin D, 9 mg of vinblastine, and 94 mg of cisplatin iv and 48 units of bleomycin by continuous infusion daily for 6 days. On November 20, 1978, the bleomycin infusion was stopped because of severe retrosternal chest pain. An electrocardiogram was initially normal, but on the following day it showed diffuse ST segment elevation consistent with acute pericarditis. A friction rub was not heard. SGOT, LDH, and cardiac fraction of CPK were normal. An echocardiogram failed to reveal pericardial effusion. Pain subsided promptly, but serial electrocardiograms revealed changes consistent with evolving pericarditis. A perfusion lung scan was normal. A collagen vascular workup and viral studies were not done. Flurazepam (Dalmane) and thiethylperazine (Torecan) were used as antiemetics and the patient was not receiving any other medication.

The patient was re-admitted on December 27, 1978, for a second course of chemotherapy. An electrocardiogram showed minor nonspecific ST-T changes. The serum HCG-beta level was normal. On December 29, 24 hours after receiving cyclophosphamide, actinomycin D, vinblastine, cisplatin, and bleomycin, he noted mild chest pain. An electrocardiogram was unchanged. On January 2, 1979, chest pain reappeared with fever and a generalized toxic eruption of the skin. An electrocardiogram again showed ST-T changes compatible with acute pericarditis. The bleomycin infusion was stopped. The patient refused to stay in the hospital, but was admitted to another hospital in March 1979 with chest pain. An abnormal electrocardiogram again suggested acute pericarditis or possibly myocardial ischemia and he was discharged to his home on aspirin and Nitro-Bid.

In May 1979, a followup electrocardiogram revealed significant change with evidence of transmural anteroseptal myocardial infarction. A thallium scan showed an area of ischemia. A treadmill stress test showed significant ST-T changes with the development of ventricular tachycardia. On May 15, 1979, cardiac catheterization and coronary arteriography were performed. The angiogram showed a moderate amount of hypokinesia along the midsegment of the anterior wall. The anterior descending coronary artery was totally occluded just distal to the origin of the diagonal artery. The diagonal artery had a 40% lesion near its origin, and minor disease was noted in the circumflex and the dominant right coronary artery. The patient is being treated with Inderal, Nitro-Bid, and aspirin and has done fairly well. He received no further chemotherapy and his beta-HCG (RIA) has been within normal limits.

### Patient 2

Patient 2, a 25-year-old white male, was admitted to the hospital on January 29, 1979, for adjuvant chemotherapy following a left radical orchiectomy and retroperitoneal lymph node dissection. The retroperitoneal tumor contained metastatic teratocarcinoma and choriocarcinoma. The serum HCG-beta level was normal following lymphadenectomy. Adjuvant chemotherapy was started and he received 1.1 g of cyclophosphamide, 1.75 mg of actinomycin D, 7.5 mg of vinblastine, and 30 units of bleomycin and 27 units/day of bleomycin by continuous infusion for 6 days followed by 220 mg of cisplatin iv on the seventh day. There were no sequelae.

The patient received a second course of chemotherapy on February 25, 1979. On February 27, 1979, during bleomycin infusion, he developed severe substernal pain aggravated by inspiration and a low-grade fever. There was no friction rub. An electrocardiogram was normal. The tracing on the following day, however, showed an ST segment elevation consistent with acute pericarditis. The infusion was stopped. SGOT, LDH, and CPK levels were normal. Serial electrocardiograms showed evolving pericarditis. Single and convalescent serum titers for common viral pathogens were normal. An echocardiogram failed to reveal a pericardial effusion. Chest pain subsided gradually over the next few days. An electrocardiogram 6 weeks later was within normal limits.

The patient received no further chemotherapy. He continues to be asymptomatic with a normal electrocardiogram and beta-HCG (RIA).

## DISCUSSION

Although acute pericarditis is usually idiopathic, cases have been observed after acute upper respiratory infection, and after trauma to the sternum and pericardium. Swarbrick and Gray (4) observed pericarditis as a presenting feature in systemic lupus erythematosus during treatment with procainamide in four patients. Life-threatening pericarditis was reported by Durkin et al (3) during treatment of two patients with lymphoma with im bleomycin. Both patients had received prior mediastinal irradiation 8 and 18 months, respectively, prior to

the appearance of pericarditis.[1] Corder and Flannery (1) and Klein (2) each reported one case of acute pericarditis following administration of actinomycin D alone or in combination with cyclophosphamide. Both patients had received prior mediastinal irradiation 15 and 21 months, respectively, prior to the appearance of pericarditis. In both cases it was assumed that actinomycin D induced a previous subclinical radiation pericarditis. Appelbaum et al (5) have observed lethal myopericarditis in four of 15 patients receiving high-dose cyclophosphamide in combination with BCNU, cytosine arabinoside, and 6-thioguanine. Harrison and Sanders (6) reported pericarditis in a patient with early daunorubicin cardiomyopathy who died.

Neither patient 1 nor patient 2 had received prior mediastinal irradiation. It is possible, of course, that the two incidents are coincidental and unrelated to chemotherapy. We have used the same drug regimen 28 times in 14 patients and have observed pericarditis in only two patients, and we are not aware of similar instances in patients treated elsewhere with this regimen.

Edwards et al (7) recently reported severe coronary arteriosclerosis in autopsy findings of two young patients treated for germ cell tumor of the testis with cisplatin, vinblastine, and bleomycin. There is a possible association between long-term treatment with these drugs and accelerated arteriosclerosis of the coronary arteries.

## REFERENCES


1. Corder MP, and Flannery EP. Possible radiation pericarditis precipitated by actinomycin D. Oncology 30:81–84, 1974.
2. Klein K. Complications of testicular tumor therapy. Int J Radiat Oncol Biol Phys 2:1049–1051, 1977.
3. Durkin WJ, Pugh RP, Solomon J, et al. Treatment of advanced lymphomas with bleomycin (NSC-125066). Oncology 33:140–145, 1976.
4. Swarbrick ET, and Gray IR. Systemic lupus erythematosus during treatment with procainamide. Br Heart J 34:284–288, 1972.
5. Appelbaum FR, Strauchen JA, Graw RG, Jr, et al. Acute lethal carditis caused by high-dose combination chemotherapy. A unique clinical and pathological entity. Lancet 1(7950):58–62, 1976.
6. Harrison DT, and Sanders LA. Letter: Pericarditis in a case of early daunorubicin cardiomyopathy. Ann Intern Med 85:339–341, 1976.
7. Edwards GS, Lane M, and Smith FE. Letter: Long-term treatment with *cis*-dichlorodiammineplatinum(II)-vinblastine-bleomycin: possible association with severe coronary artery disease. Cancer Treat Rep 63:551–552, 1979.



Misbahuddin Ahmed
Robert E. Slayton

Rush-Presbyterian-St. Luke's Medical Center
1725 W Harrison St
Chicago, IL 60612



October 15, 1979


[1]Bateman JR. Personal communication, 1979.

## Respiratory Dyskinesia Associated With Hexamethylmelamine

Respiratory dyskinesia presenting as respiratory alkalosis occurred in association with hexamethylmelamine and prochlorperazine.

## CASE REPORT

A 70-year-old white female with a 2-year history of stage III ovarian adenocarcinoma was admitted to the hospital with a possibility of sepsis due to low-grade fever, rigor, and apparent respiratory distress. Initial blood gas studies showed the following values: pH 7.76, $pO_2$ 90, and $pCO_2$ 19. Serum bicarbonate was 25 mEq/liter. Repeat arterial gas studies showed levels of 7.65 and 7.51. Sepsis was ruled out. There was no evidence of salicylate ingestion or liver, thyroid, or renal dysfunction. Chest roentgenogram and electrocardiogram were normal. The patient denied pain or dyspnea. Her deep, awkward breathing was worse when she was stressed. She had no history of psychiatric illness or previous episodes of chest pain, dyspnea, perioral numbness, or bilateral hand numbness. She had not used neuroleptic tranquilizers in the past.

Ten months prior to admission the patient started receiving hexamethylmelamine at a dose of 300 mg/m$^2$ which led to a clinical response. Pyridoxine was started to limit neurotoxicity and prochlorperazine at a dose of 15 mg twice a day was used for nausea. The dose of hexamethylmelamine was adjusted for neurotoxic side effects including limb heaviness, right arm pain, chin numbness, and scapular pain. These separate symptoms resolved with decreased doses of hexamethylmelamine. After 2 months she received minimal prochlorperazine as needed. Five months prior to admission and 5 months after prochlorperazine and hexamethylmelamine she developed several "nervous habits," eg, teeth grinding, talking with her mouth closed, and repetitive movements of her right toe and foot.

This patient was a thin, elderly, anxious woman who denied pain. Her respiration was regular and deep at 20 beats/minute, associated with soft grunting. Orobuccal dyskinesia (noted by a pursing and unpursing of the lips), minimal vermiform movements of the tongue, and choreatic movements of the right big toe and foot were present. The amplitude of toe motion could voluntarily be decreased but not eliminated. Abnormal breathing was absent during sleep.

The remainder of the patient's physical examination was remarkable only for increased abdominal girth. A mental examination revealed good memory, appropriate affect, and normal handwriting. There was no thought disorder, anxiety, pill-rolling tremor, gegenhalten, frontal release signs, or ataxia. Cranial nerves and sensation were normal. Except for absent ankle jerk, reflexes were normal.

Respiratory dyskinesia, an involuntary pattern of breathing associated with grunts and gasps, is a recently described feature of choreiform movement disorders which may lead to intermittent hyperventilation (1). Since choreiform movements increase with stress, decrease with sedation, and disappear with sleep, the patient's hyperventilation follows this pattern.

The neurotoxicity of hexamethylmelamine is well-documented (2). One case of Parkinson's-like disease (3) and two cases of tremor (4,5) have been described. Both of these movement disorders stopped when hexamethylmelamine was discontinued.

Neuroleptics have frequently been associated with tardive dyskinesia. In this case, the patient had been given prochlorperazine. Tardive dyskinesia usually occurs, however, with a higher cumulative dose and longer duration (usually > 6 months) of neuroleptics. One case of tardive dyskinesia experienced with prochlorperazine, after 4 years of use, has been reported (6). Respiratory dyskinesia has never been reported with prochlorperazine.

The combination of hexamethylmelamine and phenothiazine antiemetics may lead to respiratory dyskinesia. Tardive dyskinesia, by definition, persists after the offending agent has been stopped, and may worsen when dopamine antagonists are discontinued.

To minimize these disorders, the earliest signs of dyskinesia, especially vermiform movements of the tongue and involuntary orobuccal movements, should be recognized and the drugs should be re-evaluated.

## REFERENCES


1. Weiner JW, Goetz CG, Nausieda PA, et al. Respiratory dyskinesias: extrapyramidal dysfunction and dyspnea. Ann Intern Med 88:327–331, 1978.
2. Legha SS, Slavik M, and Carter SK. Hexamethylmelamine. An evaluation of its role in the therapy of cancer. Cancer 38:27–35, 1976.
3. Wilson WL, Bisel HF, Cole D, et al. Prolonged low-dosage administration of hexamethylmelamine (NC 13875). Cancer 25:568–570, 1970.
4. Denefrio JM, and Vogel CL. Phase II study of hexamethylmelamine in women with advanced breast cancer refractory to standard cytotoxic therapy. Cancer Treat Rep 62:173–175, 1978.
5. Smith JP, and Rutledge FN. Random study of hexamethylmelamine, 5-fluorouracil and melphalan in treatment of advanced carcinoma of the ovary. Symposium on Ovarian Carcinoma. Natl Cancer Inst Monogr 42:169–172, 1975.
6. Klawans HL, Bergen D, and Bruyn GW. Neuroleptic induced tardive dyskinesia in nonpsychotic patients. Arch Neurol 30:27–35, 1975.



George B. Murray
Donna B. Greenberg[1]


[1]We are indebted to Drs. Robert W. Carey and W. Davies Sohier, Jr, oncologists, for bringing this case to our attention and for criticizing the manuscript.


Department of Psychiatry
Massachusetts General Hospital
Boston, MA 02114



October 22, 1979


## Acute Ventricular Fibrillation and Death During Infusion of 4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA)

4′-(9-Acridinylamino)methanesulfon-*m*-anisidide (AMSA) is an acridine which has recently completed phase I clinical trials (1–5). Phase I and early phase II studies have documented hints of antitumor activity of AMSA in acute myelogenous leukemia (6), lymphoma (1), breast cancer (6), and melanoma (7).

The following is a case report of ventricular fibrillation which occurred during an infusion of AMSA.

## CASE REPORT

Patient 1, a 42-year-old female, had acute myelogenous leukemia diagnosed in June 1977. Over a 2-year period she had received rubidazone (total dose, 800 mg/m$^2$), vincristine, cytosine arabinoside, and prednisone induction therapy followed by vincristine, cytosine arabinoside, and prednisone maintenance therapy. She responded to this therapy with a complete remission until January 1979. She failed attempted re-induction therapy with Adriamycin plus vincristine, cytosine arabinoside, and prednisone. From March 15, 1979, to April 8, 1979, she received three courses of methyl-GAG. She did not respond to methyl-GAG and her treatment was changed to VP-16-213 at a dose of 240 mg/m$^2$ daily for 5 days. She had no side effects from this high dose of VP-16-213 and had a brief partial response. By June 14, 1979, her leukemia worsened with a wbc count of 5900/mm$^3$ (62% blast cells) and a platelet count of 10,000/mm$^3$. She was febrile to 101°C.

In an attempt to bring her leukemia under control, and after informed consent was given, she started receiving AMSA on June 14, 1979. The dose given was 35 mg/m$^2$ daily × 7 via a peripheral vein infused in 500 ml of 5% dextrose in water over 1 hour. This dose represented a downward dose adjustment from the usual 75–90 mg/m$^2$ daily × 7 because her bilirubin level was 6.8 mg/100 ml, her alkaline phosphatase level was 247 IU/ml, and her SGOT level was 429 IU/ml. AMSA blood levels are known to be prolonged in patients with hepatic dysfunction (8). The liver function abnormalities were thought to represent a post-transfusion hepatitis or leukemic infiltration of the liver.

The first course of AMSA was well-tolerated. By Day 6 of the course the patient's wbc count was 1100/mm$^3$ with

0 blast cells and her platelet count was 1000/mm$^3$. Bone marrow aspirate showed hypocellularity with 20% blast cells. The wbc count remained at 1000/mm$^3$ over the next 18 days. Platelet transfusions were given daily until July 13, 1979, when her platelet count was 57,000/mm$^3$. Her serum bilirubin and SGOT levels had also returned to normal by July 13. However, a bone marrow examination on that day revealed increased cellularity with 70% blast cells. The peripheral wbc count was 1200/mm$^3$ with 20% blast cells. Because of worsening leukemia the patient was scheduled for a second course of AMSA. Before the start of the second course of AMSA the patient's serum potassium level was 2.3 mEq/liter. With more vigorous replacement (150 mEq/day iv and 120 mEq/day orally) her serum potassium level rose to 2.7 mEq/liter after 3 days of therapy. Serum calcium concentration at the start of therapy was 8.5 mg/100 ml. Baseline electrocardiogram was normal. Treatment with the second course of AMSA was begun on July 16 at a dose of 75 mg/m$^2$ daily × 7 since her liver function tests were normal. The drug was administered through a central venous line (catheter tip in superior vena cava) in 500 ml of 5% dextrose in water over 1 hour. The patient tolerated the first 3 days of treatment without incident. However, on the fourth day of treatment, after about 250 ml of AMSA had been infused, the patient became pulseless and unresponsive. Her breathing continued. Electrocardiogram showed ventricular fibrillation. Direct current countershock was applied and resuscitation was attempted but the patient did not respond to these measures and died.

## DISCUSSION

This case is reported to call attention to the possible relationship between the administration of AMSA and the development of a fatal ventricular arrhythmia. The etiology of the arrhythmia is not entirely clear. The patient could have had an arrhythmia on the basis of a structural lesion of the heart (leukemia, infection, or atherosclerosis). Unfortunately, autopsy permission was not granted to rule out these possibilities. The hypokalemia alone could also have caused the arrhythmia, although the patient did have a normal electrocardiogram even with a potassium level of 2.3 mEq/liter. The central venous catheter had been in place several days and had not changed position to irritate the ventricle. In addition, AMSA had been infused into the same line for 3 days before the arrhythmia and had not caused a problem. The patient had had prior anthracyclines but there is currently no evidence that cardiac arrhythmias are late side effects of anthracyclines. The occurrence of the ventricular arrhythmia during the AMSA infusion strongly suggests a relationship between the drug administration and the infusion of AMSA. In dogs and monkeys the drug did not produce any cardiotoxicity so it is questionable whether the compound is a direct cardiac toxin.

Besides a direct cardiotoxic effect of AMSA itself, it could be postulated that the vehicle for reconstitution of AMSA (ie, *N*,*N*-dimethylacetamide and lactic acid) may be cardiotoxic. Dimethylacetamide had been used as a vehicle for other drugs without incident. A synergistic cardiotoxic effect between hypokalemia and AMSA is also a possibility. This is particularly likely since we have recently become aware of ventricular arrhythmias noted with AMSA infusions in two patients with electrolyte imbalance at the M. D. Anderson Hospital and Tumor Institute (Houston, TX) (9).

In summary, we report here a case of fatal ventricular fibrillation which occurred during an infusion of AMSA. At this point it would seem prudent to attempt to correct electrolyte abnormalities in all patients before AMSA is administered. In addition, the patients receiving AMSA should be watched closely during the infusion of the drug, with cardiac monitoring and resuscitation facilities close at hand. This vigilance should continue until the precise etiology of the drug-associated arrhythmias is ascertained.

## REFERENCES


1. Von Hoff DD, Howser D, Gormley P, et al. Phase I study of methanesulfonamide, *N*-[4-(9-acridinylamino)-3-methoxyphenyl]-(*m*-AMSA) using a single-dose schedule. Cancer Treat Rep 62:1421–1426, 1978.
2. Legha SS, Gutterman JU, Hall SW, et al. Phase I clinical investigation of 4′-(9-acridinylamino)methanesulfon-*m*-anisidide (NSC 249992), a new acridine derivative. Cancer Res 38:3712–3716, 1978.
3. Goldsmith MA, Bhardwaj S, Ohnuma T, et al. Phase I study of *m*-AMSA in patients with solid tumors and leukemias. Proc Am Assoc Cancer Res and ASCO 20:344, 1979.
4. Schneider R, Sklaroff R, Ochoa M, Jr, et al. Phase I trial of AMSA (4′-(9-acridinylamino)-methanesulfon-*m*-anisidide). Proc Am Assoc Cancer Res and ASCO 20:114, 1979.
5. Van Echo DA, Gormley P, Lichtenfeld JL, et al. A phase I study of *m*-AMSA. Proc Am Assoc Cancer Res and ASCO 20:422, 1979.
6. Legha SS, Bodey GP, Keating MJ, et al. Early clinical evaluation of acridinylamino-methanesulfon-*m*-anisidide (AMSA) in patients with advanced breast cancer and acute leukemia. Proc Am Assoc Cancer Res and ASCO 20:416, 1979.
7. Hall SW, Benjamin RS, Legha SS, et al. AMSA: a new acridine derivative with activity against metastatic melanoma. Proc Am Assoc Cancer Res and ASCO 20:372, 1974.
8. Hall SW, Benjamin RS, Legha SS, et al. Clinical pharmacokinetics of the new antitumor agent AMSA. Proc Am Assoc Cancer Res and ASCO 20:175, 1979.
9. Legha SS, Latreille J, McCredie KB, et al. Neurologic and cardiac rhythm abnormalities associated with 4′-(9-acridinylamino)methanesulfon-*m*-anisidide (AMSA) therapy. Cancer Treat Rep 63:2001–2003,1979.



Daniel D. Von Hoff*
David Elson
Gary Polk
Charles Coltman, Jr

*To whom reprint requests should be addressed.

Department of Medicine
Division of Oncology
University of Texas Health Science Center at San Antonio
7703 Floyd Curl Dr
San Antonio, TX 78284

October 30, 1979

## Multiple Ventricular Extrasystoles Following Administration of 4′-(9-Acridinylamino)-methanesulfon-*m*-anisidide (AMSA)[1]

To supplement reports by S. S. Legha from the M. D. Anderson Hospital and Tumor Institute (Houston, TX) and D. D. Von Hoff from the University of Texas Health Science Center (San Antonio), we would like to call attention to two further patients with cardiotoxicity receiving 4′-(9-acridinylamino)methanesulfon-*m*-anisidide (AMSA). Both of our patients were receiving AMSA as treatment of primary liver cancer in an Eastern Cooperative Oncology Group pilot trial. Both patients were treated with AMSA at a dose level of 120 mg/m$^2$ every 28 days.

### Patient 1

Patient 1 was a 50-year-old man. Following the second dose of AMSA (Day 28 of treatment), the patient collapsed several minutes after the infusion had been completed. Resuscitation was attempted but failed and the cause of death was presumed to be acute cardiotoxicity.

### Patient 2

Patient 2 was a 38-year-old man. Immediately following the administration of the first dose of AMSA the patient complained of not feeling well. An electrocardiogram revealed multiple ventricular extrasystoles. Figure 1 shows the appearance of the electrocardiogram. These extrasystoles disappeared within ½ hour. The following day the patient developed hemoperitoneum and died shortly after needle aspiration of the abdomen had confirmed that this had occurred.

Neither of the patients had clinical evidence of previous cardiac disease or had had prior anthracycline exposure. Unfortunately, electrocardiographic examination had not been performed on either patient prior to treatment.

G. Falkson

Department of Cancer Chemotherapy
H. F. Verwoerd Hospital and the
University of Pretoria
Private Bag X169
Pretoria 001, Republic of
South Africa

Oct 31, 1979

*Supported by Public Health Service grant CA-21692 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

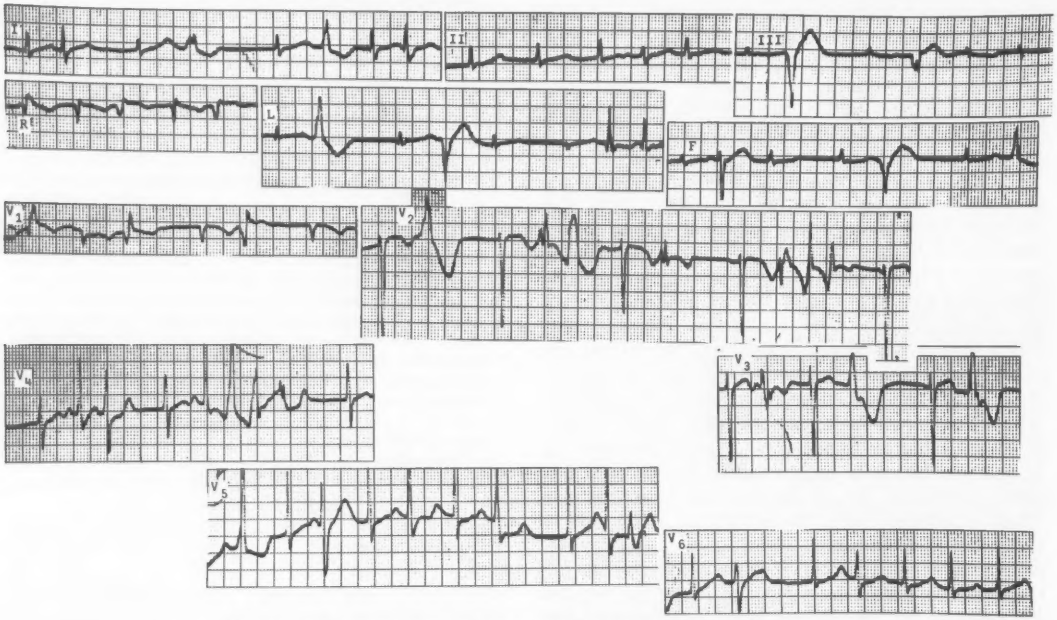



FIGURE 1.—Electrocardiogram showing multiple ventricular extrasystoles.

## Vincristine-Induced Orthostatic Hypotension: A Prospective Clinical Study

Vincristine has been reported to cause orthostatic hypotension (1–4), although some question has been raised as to whether this reaction is due to the vincristine (4) or to other agents used in combination. Since patients with malignancies often experience symptoms which could be due to mild or subclinical orthostatic hypotension, a prospective study was initiated to detect subtle forms of this reaction in patients receiving vincristine.

Twenty-six patients receiving vincristine in combination with other agents were studied (ten females and 16 males ranging in age from 19 to 64 years). Of these 26 patients, 23 were receiving chemotherapy for lymphomas or leukemias and three were receiving chemotherapy for nonhematologic malignancies. None of these patients had known neurologic disorders, brain or spinal cord involvement by tumor.

Orthostatic hypotension was documented by placing the patient on a tilt table and obtaining three baseline blood pressures and pulses at approximately 1-minute intervals. The tilt table was then elevated to the vertical position so that the patient's weight was supported by a cushioned saddle and the patient's feet did not touch the floor. Blood pressures and pulse rates were then obtained for an additional 5 minutes.

The cumulative dose of vincristine received by the patients studied varied from 6 to 56.8 mg, with a median dose of 20 mg administered over the preceding 12 months. Tilt-table observations were made within 30 minutes of vincristine administration in all but two patients who were studied 90 and 45 minutes after their last dose.

Orthostatic hypotension was observed in only one patient, a 22-year-old white male receiving MOPP therapy (mechlorethamine, vincristine, prednisone, and procarbazine) for Hodgkin's disease. The only other medication this patient received with the vincristine was mechlorethamine. Prednisone and procarbazine were not started until the next morning. He had previously experienced symptoms of orthostatic hypotension after the third dose of vincristine at a cumulative dose of 6 mg. He described marked weakness and near syncopal episodes with assumption of the upright posture, lasting 3–6 hours after each of his last four doses of vincristine. His baseline systolic pressure was 93 mm Hg with a pulse rate of 82 beats/minute. When elevated to the vertical position, his systolic pressure became inaudible and palpable at approximately 40 mm Hg after 5 minutes. After 7 minutes, the blood pressure became audible again at 88 mm Hg. The pulse rate increased to approximately 100 beats/minute. With the assumption of the erect position, the patient complained of marked dizziness and became diaphoretic. This tilt test was repeated 5 minutes later with similar results. However, when the test was repeated the following morning (approximately 16 hours later), there were no symptoms or findings of orthostatic hypotension. Vincristine was eliminated from this patient's subsequent chemotherapy courses, which were administered at another hospital. Although blood pressure values were not obtained when he received these subsequent doses of mechlorethamine alone, he did not experience the same severe symptoms after injection.

None of the other patients studied had orthostatic hypotension. The mean systolic pressure remained essentially unchanged in five patients, decreased by $< 10$ mm Hg in seven patients, and increased by 4–20 mm Hg in 13 patients. All but two of the patients experienced an increase in pulse rate varying from 1 to 58 beats/minute, with a median increase in pulse rate of 12 beats/minute. Seven of the 26 patients were experiencing sensory or motor neurotoxic effects due to the vincristine, including mild paresthesias and distal motor weakness in the patient who had orthostatic hypotension.

In this small group of patients, no instances of subclinical orthostatic hypotension were discovered. Yet, more than one half of these patients (14 of 26) had received cumulative doses of $\geq 20$ mg of vincristine. In fact, the only patient to experience this side effect began to do so only after a total of 6 mg, suggesting that it is not dose-related.

The mechanism for this hypotensive effect is not clear and the present study was unable to clarify this point. The fact that our patient was experiencing mild paresthesias associated with the hypotension suggests that it represents another neurotoxic reaction. A report by Carmichael et al (2) suggests that an impairment in norepinephrine excretion may be the mechanism for this reaction. They demonstrated that norepinephrine excretion was markedly depressed following vincristine administration in a patient who experienced orthostatic hypotension. One difficulty with interpreting this study is that this patient with reticulum cell sarcoma had central nervous system involvement due to lymphomatous infiltration of the fourth ventricle, which could have affected his baroreceptor response. However, they observed decreased norepinephrine urinary excretion in two other patients receiving vincristine, neither of whom experienced any orthostatic symptoms. It was also demonstrated that vincristine does not appear to interfere with the uptake or storage of norepinephrine in tissues. If these observations are correct, then the most likely mechanism is a lesion in the postganglionic nerve induced by vincristine and interfering with the release of norepinephrine from the nerve endings.

In conclusion, orthostatic hypotension appears to be a relatively uncommon neurologic complication of vincristine therapy occurring in 4% (one of 26) of patients studied prospectively. No instances of subtle or subclinical hypotension were found.

## REFERENCES


1. AISNER J, WEISS HD, CHANG P, ET AL. Orthostatic hypotension during combination chemotherapy with vincristine (NSC-67574). Cancer Chemother Rep 58:927–930, 1974.
2. CARMICHAEL SM, EAGLETON L, AYERS CR, ET AL. Orthostatic hypotension during vincristine therapy. Arch Intern Med 126:290–293, 1970.
3. MCLEOD JA, and PENNY R. Vincristine neuropathy: an electrophysiological and histological study. J Neurol Neurosurg Psychiatry 32:297–304, 1969.
4. ROSENTHAL S, and KAUFMAN S. Vincristine neurotoxicity. Ann Intern Med 80:733–737, 1974.



Nicholas J. DiBella

Hematology-Oncology Service
Fitzsimons Army Medical Center
Aurora, CO 80045


November 13, 1970

## Phenytoin Sodium-Induced Alterations in the Pharmacokinetics of Misonidazole in the Dog

Misonidazole (MIS) is presently undergoing extensive clinical trials as a hypoxic cell radiosensitizer. Many patients receiving MIS may also require sedatives and anticonvulsants, such as phenobarbitone and phenytoin, which are potent inducers of liver microsomal enzymes, as part of their general medical care. We report in this paper the effects of pretreatment with phenytoin on the pharmacokinetics of MIS in the dog.

Two adult, male, and crossbred dogs weighing 17 and 20 kg were used in this study. Their routine hematologic and biochemical parameters were found to be normal before and during the study. MIS was administered to both dogs at a dose of 150 mg/kg by iv bolus injections (60 ml in 30 seconds). Blood samples were then removed over a 36-hour period. Following a period of 7 days, phenytoin (Parke, Davis and Co, Detroit, MI) was administered daily at the recommended veterinary dose of 17 mg/kg orally. The dose was divided into equal portions for twice daily administration over a 12-day period. Three days after the withdrawal of phenytoin, MIS was again administered to the dogs and blood samples were taken as before. Plasma samples were stored at –20°C and assayed by high-performance liquid chromatography for MIS and its *O*-demethylated metabolite, desmethylmisonidazole (1).

The pharmacokinetic data are presented in table 1. Values for the kinetic parameters recorded before phenytoin pretreatment were found to be similar to those previously recorded in normal dogs (2,3). The differences in the pharmacokinetic data between the first and second MIS dose can be attributed to the effects of phenytoin pretreatment since repeated MIS administration does not alter the drug's pharmacokinetic behavior (2). The results are summarized as follows.

*A.* Apparent peak plasma concentrations were reduced only slightly (3% and 10%) and were recorded either at the earliest time point (5 minutes) or, in one case, at 30 minutes. This is consistent with our earlier studies in which abnormal disposition kinetics were noted during the distribution phase (2).

*B.* The plasma half-life was reduced by 38% and 39% in the two dogs. The area under the curve (AUC) of concentration versus time plot ($AUC_{0-\infty}$) was decreased by 33% and 40% and the clearance (dose/AUC) was increased by the same amount.

*C.* Peak plasma desmethylmisonidazole concentrations were increased by 33% and 66% but peak times were unaltered. The AUC values for desmethylmisonidazole were increased by 10% and 21% but the AUC values for total 2-nitroimidazole (MIS and desmethylmisonidazole) were reduced by 26% and 37%.

These data indicate that the reduced plasma half-life and AUC values for MIS are due, at least in part, to increased *O*-demethylation of MIS to desmethylmisoni-

TABLE 1.—Pharmacokinetic data

| Treatment | Misonidazole | | | | Desmethylmisonidazole | | | |
|---|---|---|---|---|---|---|---|---|
| | Peak concentration (μg/ml) | Time of peak concentration (mins) | Plasma half-life* (hrs) | AUC (μg·h/ml) | Plasma clearance (l/kg/h) | Peak concentration (μg/ml) | Time of peak concentration (hrs) | AUC (μg·h/ml) |
| *Dog 1* | | | | | | | | |
| No pretreatment | 221 | 5 | 5.0 (4.8–8.2) | 1396 | 0.11 | 12 | 5 | 202 |
| Phenytoin pretreatment | 200 | 30 | 3.1 (2.9–3.2) | 939 | 0.16 | 16 | 5 | 244 |
| *Dog 2* | | | | | | | | |
| No pretreatment | 222 | 5 | 7.5 (6.9–8.3) | 2566 | 0.059 | 9 | 3 | 187 |
| Phenytoin pretreatment | 215 | 5 | 4.6 (4.4–4.9) | 1540 | 0.097 | 15 | 3 | 206 |

*95% confidence limits for half-life in parentheses.

dazole as a result of the induction of hepatic mixed-function oxidases by phenytoin.

Similar effects on MIS metabolism have been observed with phenobarbitone and phenytoin pretreatment in the mouse (4), and this agreement lends support to the possibility that this may also occur in man. If toxicity is related to MIS exposure (AUC), it is possible that it may be reduced by more rapid drug clearance.

## REFERENCES


1. WORKMAN P, LITTLE CJ, MARTEN TR, ET AL. Estimation of the hypoxic cell-sensitiser misonidazole and its *O*-demethylated metabolite in biological materials by reversed-phase high-performance liquid chromatography. J Chromatogr 145:507–512, 1978.
2. WHITE RAS, WORKMAN P, FREEDMAN LS, ET AL. The pharmacokinetics of misonidazole in the dog. Eur J Cancer 15:1233–1242, 1979.
3. WHITE RAS, WORKMAN P, OWEN LN, ET AL. The penetration of misonidazole into spontaneous canine tumours. Br J Cancer 40:284–294, 1979.
4. WORKMAN P. Effects of pretreatment with phenobarbitone and phenytoin on the pharmacokinetics and toxicity of misonidazole in mice. Br J Cancer 40:335–353, 1979.



R. A. S. White*

Department of Clinical
Veterinary Medicine
University of Cambridge
Madingley Rd
Cambridge, CB3 OES
England

P. Workman

M. R. C. Clinical Oncology
and Radiotherapy Unit
University of Cambridge
Hills Rd
Cambridge, CB2 2QH
England



November 13, 1979



*To whom reprint requests should be sent.


## Avascular Necrosis of the Femoral Head: Association With Adjuvant Chemotherapy for Breast Carcinoma

Two patients treated in adjuvant fashion for breast carcinoma with combination chemotherapy including corticosteroids developed avascular necrosis of bone. Although it is a rare, but well-known, consequence of prolonged high-dose corticosteroid therapy, avascular necrosis has not previously been reported in such patients.

## CASE REPORTS

### Patient 1

Patient 1, a 62-year-old female with a negative past medical history and a breast tumor of 3 months' recognized duration, underwent modified radical mastectomy for infiltrating duct cell carcinoma in June 1977. Twelve of 18 axillary lymph nodes contained metastatic carcinoma. She was advised to take adjuvant chemotherapy. With informed consent, she was randomized to receive treatment in Cancer and Leukemia Group B protocol 7581, a study program comparing three different chemotherapeutic regimens. Pretreatment chest roentgenogram, bone survey, and bone scan were within normal limits. Chemotherapy was started in June 1977 and consisted of 6 weeks' induction treatment with cyclophosphamide, methotrexate, 5-fluorouracil (5-FU), vincristine, and prednisone (CMFVP) followed by monthly cycles of the same drugs.

During induction therapy, drug doses were as follows: 80 mg/m$^2$ orally of cyclophosphamide on Days 1–42; 40 mg/m$^2$ iv of methotrexate weekly; 500 mg/m$^2$ iv of 5-FU weekly; 1.4 mg/m$^2$ iv of vincristine weekly; and 40 mg/m$^2$ orally of prednisone on Days 1–21, tapered and discontinued by Day 28. During maintenance therapy, iv drugs were given at identical doses on Days 1 and 8 of each 28-day cycle. The oral cyclophosphamide dose was increased to 100 mg/m$^2$ on Days 1–14 and the prednisone dose was given at 40 mg/m$^2$/day on Days 1–14 and then discontinued without tapering.

Bone scan and complete skeletal roentgenograms from December 1977 were within normal limits. In March 1978, the patient began to complain of right hip pain. In April 1978, bone survey and tomograms of the right hip revealed aseptic necrosis of the right femoral head. Bone scan revealed increased uptake in that region alone. Because of increasing disability, total hip replacement was performed in June 1978. Pathologic examination of the excised femoral head revealed avascular necrosis. A total prednisone dose of 1235 mg had been given over the previous 10 months. The patient made a full recovery and remains without evidence of recurrent neoplastic disease or contralateral hip disease 24 months after mastectomy.

### Patient 2

Patient 2, a 47-year-old female with a negative past medical history, had stage III ($T4_BN1_BM_O$) carcinoma of the right breast of 53 months' recognized duration in March 1975. Pretreatment chest roentgenogram, bone scan, and bone survey were within normal limits. She received five monthly cycles of cyclophosphamide, methotrexate, and 5-FU followed by a modified radical mas-

tectomy. Postoperatively, she received monthly cycles of the same drugs for 2 years. Drug doses were identical to those given to patient 1 during maintenance chemotherapy. In December 1977, 3 months after chemotherapy was discontinued, an isolated right axillary recurrence was proven by excisional biopsy. Subsequently, with informed consent, she started to receive an uninterrupted regimen of chemotherapy with CMFVP which was given in three 6-week intensive cycles at ten weekly intervals according to a local pilot study. Drug doses were identical to those of the induction chemotherapy given to patient 1. This treatment was initiated in February 1978 and was completed in October 1978. In April 1979, she complained of left hip pain. Bone scan revealed increased uptake in the left femoral head. Subsequent bone roentgenograms and tomograms revealed bilateral flattening, cortical sclerosis, and subchondral lucencies compatible with bilateral avascular necrosis of the femoral heads. A total prednisone dose of 1465 mg had been given over 38 weeks for 6 months prior to this diagnosis. Physical examination revealed no evidence of recurrent disease, and the patient has been without macroscopic evidence of breast cancer for 18 months as of this report.

## DISCUSSION

The pathogenesis of avascular necrosis of bone remains obscure. Nevertheless, this complication of the corticosteroid treatment of a variety of diseases, including many in which this lesion is not part of the natural history of the underlying disease, is an example of iatrogenic morbidity. More than 175 cases of corticosteroid-induced avascular necrosis were reported by 1976 (1–3). Common features included a lack of association with significant trauma; corticosteroid administration in excess of physiologic requirements for long periods, regularly or intermittently; and primary disease entities not known to be related to the development of atraumatic avascular necrosis. Avascular necrosis has been observed in eight patients with lymphomas treated by corticosteroid-containing regimens (1,2). Biopsy-proven avascular necrosis of the humeral head has been reported in a single patient with stage II breast carcinoma, but she had had postoperative radiotherapy followed by cyclophosphamide, methotrexate, and 5-FU, and a cervical laminectomy for a herniated intervertebral disc followed by prolonged arm immobilization (4).

The present report establishes by pathologic evidence in one patient and by classic clinical and radiologic findings in the other the diagnosis of avascular necrosis. It leads to a new awareness of the possible association between corticosteroid-induced avascular necrosis and surgical adjuvant chemotherapy. Two of 40 patients with stage II or III breast carcinoma at this institution treated with corticosteroid-containing regimens have developed avascular necrosis. However, the efficacy of CMFVP in patients with stage II breast carcinoma suggests that such morbidity is entirely acceptable compared to the alternative of inexorable death following clinically evident metastatic disease (5). The likelihood that the advent of postsurgical adjuvant chemotherapy of breast carcinoma will have a major impact on survival in the near future is clear; also, it is evident that previously unreported drug toxic effects will be noted.

## REFERENCES


1. Ihde DC, and DeVita VT. Osteonecrosis of the femoral head in patients with lymphoma treated with intermittent combination chemotherapy (including corticosteroids). Cancer 36:1585–1588, 1975.
2. Sweet DR, Jr, Roth DG, Kesser RK, et al. Avascular necrosis of the femoral head with combination therapy. Ann Intern Med 85:67–68, 1976.
3. Fisher DE, and Bickel WH. Corticosteroid-induced avascular necrosis. A clinical study of seventy-seven patients. J Bone Joint Surg [AM] 53:859–873, 1971.
4. Obrist R, Hartmann D, and Obrecht JP. Letter: Osteonecrosis after chemotherapy. Lancet 1(8077):1316, 1978.
5. Cooper RG, Holland JF, and Glidewell O. Adjuvant chemotherapy of breast cancer. Cancer 44:793–799, 1979.



Marjorie Perloff*
Gerson J. Lesnick

Mt. Sinai School of Medicine
New York, NY 10032



November 16, 1979



*Reprint requests should be addressed to: Dr. M. Perloff, Division of Oncology, Department of Medicine, College of Physicians and Surgeons of Columbia University, 630 W 168th St, New York, NY 10032.

# Proceedings of the Symposium on Designs for Clinical Cancer Research

Chairman: Nell Sedransk
Co-Chairman: Stephen Carter

April 13–15, 1978
New Orleans, Louisiana

# Ethics and Designs: Laetrile Trials as an Example

**Guy R. Newell* and Neil M. Ellison[1]**

## SUMMARY

**The National Cancer Institute instituted a search for cancer patients who have had objective benefit from Laetrile therapy. Cases have been solicited through a nationwide mailing to > 450,000 physicians and other health professionals, through publicity in the medical and lay press, and by contact with the pro-Laetrile groups. As a result, a retrospective analysis of the medical records of patients who volunteered their cases and who believed they had objective antitumor benefit from Laetrile therapy has been carried out. Similar criteria for response that are used for any National Cancer Institute-sponsored investigational agent were used to evaluate possible Laetrile antitumor activity. This analysis was not designed to determine a response rate to Laetrile. It is uncontrolled, only positive responders were solicited, and the total number of cases treated with Laetrile was not available. It is believed that if any objective benefit for Laetrile therapy exists, then at least hints of this should be made evident by this retrospective evaluation. Data collected were used to assist in deciding if clinical trials of Laetrile are warranted.**

[Cancer Treat Rep 64:363–365, 1980]

Before any new drug is introduced into clinical cancer trials, it must be approved by the Food and Drug Administration (FDA) in the form of an Investigational New Drug (IND) exemption. Included in the requirements for approval are scientific data confirming the substance's effectiveness in animal models and its relative lack of toxicity.

Laetrile, a cyanogenic glycoside, and similar other compounds have been used for many years by a relatively small population of physicians and patients for the prevention, palliation, or cure of cancer. This drug is often given in conjunction with metabolic therapy, a term that refers to the use of a special diet, supplementary vitamins, enzymes, or chelating agents. Most claims concerning the modes of action or effectiveness of this therapy are not substantiated by well-controlled scientific research.

The National Cancer Institute (NCI) and several nonfederal institutions have tested Laetrile extensively in their preclinical screens. The animal tests have failed to show reproducible anticancer activity. In addition, the human cases presented by the compound's advocates have not been documented well enough to demonstrate any definite benefit from the drug.[2] The NCI reviewed the animal and human data for Laetrile as it would for any other drug it evaluated and decided that further testing was not indicated. Since resources allow only a limited number of drugs to enter human trials, other compounds which demonstrated much greater anticancer activity in experimental screens were given preference.

In spite of the negative findings by the NCI (1,2) and other major cancer centers (3,4), public interest in Laetrile has not decreased. The use of Laetrile has continued to be a public issue with strong emotional overtones. Thousands of cancer patients are requesting Laetrile, and their access to it has been legalized in 17 states.[3] A

[1]Office of the Director, National Cancer Institute, Bethesda, MD.

**Reprint requests to:* Guy R. Newell, MD, Division of Cancer Prevention, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Texas Medical Center, Houston, TX 77030.

[2]Supplementary Report by the Cancer Advisory Council on the Treatment of Cancer with Beta-Cyanogenetic Glucosides. Prepared by the Cancer Advisory Council, State of California Department of Public Health, 1965.

[3]Alaska, Arizona, Delaware, Florida, Idaho, Illinois, Indiana, Kansas, Louisiana, Maryland, Nevada, New Hampshire, New Jersey, Oklahoma, Oregon, Texas, and Washington.

federal court order has enabled anyone in the US to obtain it for his or her own personal use.[4]

In response to growing public and professional interest, the NCI indicated that it would reconsider conducting a clinical trial of Laetrile. It was generally believed that a prospective clinical trial was at least feasible, but many were reluctant to circumvent the usual requirements of effectiveness and safety for the introduction of a new drug. Laetrile's efficacy remains unproven and there is mounting evidence that it may be harmful.

Toxicity of Laetrile has been reviewed by the FDA (5). Toxic side effects include cyanide poisoning resulting in death, dizziness, nausea, vomiting, hypertension, shock, stupor, coma, and respiratory failure. Reports of toxic effects from Georgetown University (Washington, DC) include allergic reactions ranging from acute episodes of fever, rash, and gastrointestinal symptoms to proximal muscular weakness (6). Samples of Laetrile products examined by the NCI were found to be chemically subpotent, microbiologically contaminated, and pyrogenic (7).

## ETHICAL CONSIDERATIONS

As the NCI was formulating its plans for further Laetrile evaluation, it became apparent that discord on this subject existed between nationally renowned clinicians. Strong ethical objections to the conduct of a prospective clinical trial were expressed by Lipsett and Fletcher (8), while Moertel (9) disagreed and argued for its immediate implementation. A discussion of the ethics involved is of great importance.

The NCI believes that physicians cannot ethically offer a drug to patients which is not proven to be safe or effective in animals or man, especially when new compounds constantly become available that have active antitumor properties in preclinical screens. The NCI realizes that the animal preclinical screens probably are not 100% reliable in predicting drug activity in human cancers. The purpose of the screens is to set priorities for compounds to enter the clinical arena. Ideally, the human trial is the best screen, but the NCI cannot test all agents in man. Laetrile, like any other drug with low screening activity, has never had an absolutely conclusive test of its effectiveness in man.

Another concern has been that the public will view any action taken by the NCI to test Laetrile as an endorsement of the drug. If a prospective clinical trial were performed, it is possible that a private individual might conclude that the NCI believed that there were sound scientific and medical indications for the use of Laetrile. The NCI also did not want to set a damaging precedent by subverting the FDA requirements for safety and efficacy. For these reasons it was felt that a clinical trial of Laetrile should only be embarked upon after careful evaluation of existing human data on the drug.

The question is posed whether any medical treatment is unethical if the patient has complete knowledge of the issue and still consents to participate. This presupposes that every individual patient has total access to all information and is qualified to act upon it. The reality is that there are many remarkable but unsubstantiated claims for Laetrile that could easily be misconstrued as fact. It is the role of the physician to critically review and determine the credibility of the data for the patient and to select a therapeutic plan that he or she believes will be most beneficial. The NCI has done this in the past in its consideration of the Laetrile problem. However, social and political pressures have mandated further review of issues relating to a Laetrile clinical trial.

## DESIGN CONSIDERATIONS

In the face of public pressure, increasing use of Laetrile, continued passage of state laws legalizing its use, and the ethical controversy, the NCI announced its intention to pursue a middle-of-the-road course in its reinvestigation of the Laetrile issue (10). We have attempted to document bona fide responses to Laetrile by conducting a retrospective review of case records. The Laetrile users make up a form of uncontrolled clinical trial which is ongoing in the community and we felt a responsibility to investigate these cases in order to document responses in people who have taken Laetrile. These data would allow the NCI to be better able to make a final judgment on whether to initiate a clinical therapeutic trial of its own. Documentation of objective beneficial responses would help considerably in resolving the ethical issue.

This decision was made after discussion with members of the academic community, the Director's Office of the National Institutes of Health, the FDA, the Center for Disease Control, and the Armed Forces Institute of Pathology. The working hypothesis was a simple one: to document any objective beneficial antitumor responses to Laetrile. Criteria for case entrance into the analysis were formulated in an attempt to allow objective evaluation of beneficial effects of Laetrile (with or without metabolic therapy) alone. In order to be evaluated, the following conditions had to be met.

(*a*) Consent of the patient or surviving next-of-kin had to be provided to allow the NCI to obtain any pertinent medical records.

(*b*) Histologic slides of the tumor had to be available for review to confirm the pathologic diagnosis of cancer.

(*c*) Measurable disease had to be present to determine response to treatment. Examples would include sequen-

[4]Rutherford vs US. Decision by Federal District Court Judge Luther Bohanon, Apr 8, 1977; Dec 5, 1977.

tial radiographs, scans, or physical measurements of tumor shrinkage.

(*d*) Adequate documentation of each individual case history was required.

(*e*) Laetrile (with or without metabolic therapy) had to have been used as the only cancer treatment for at least 30 days. Patients should not have received any of the conventional treatments (radiation, surgery, chemotherapy, immunotherapy) for at least 30 days before Laetrile treatment, but may have received these treatments at any other time.

(*f*) All written materials had to be in English.

After the data were collected, each evaluable case was individually summarized, coded, and presented to an expert medical panel consisting of non-NCI oncologists. Coded cases of Laetrile-treated patients were mixed in with similarly coded cases of patients treated with NCI-sponsored investigational anticancer drugs. Without knowing which agent was received, the panel was asked to decide for each case whether a significant anticancer response occurred. The purpose of this blind approach was to provide an unbiased assessment of the effects of Laetrile by the reviewers. The requirements for establishing a response to Laetrile treatment were precisely those necessary to establish efficacy of any NCI-tested anticancer drug. These consist of the following: *complete response*—disappearance of all evidence of tumor by physician examination, and reversion to normal of all laboratory or roentgenographic abnormalities attributable to tumor for at least 30 days; *partial response*—reduction by at least 50% of the sum of the cross-diameters of all measurable tumor for at least 30 days; and *mixed response*—reduction by at least 50% of the sum of the products of the cross-diameter of any lesion for at least 30 days, with other lesions remaining stable.

Although solicitations requesting clinical data on patients benefiting from Laetrile were sent to > 450,000 health professionals, the response was surprisingly small (11). Only 93 cases were submitted for evaluation. Of these, 26 cases were eliminated because of lack of adequate information. In the remaining 67 cases, records were examined by a panel of 12 oncologists in the "blinded" fashion described above. The panel judged that six Laetrile courses (9%) had resulted in objective antitumor response (two complete and four partial regressions). The objective regression rate of < 10% was obviously less than impressive, but did represent a hint of activity for Laetrile. It is readily apparent that this retrospective review could in no way determine a response rate for Laetrile since the denominator from which these six regressing patients were culled could be anything from 67 to > 100,000.

Several major reservations have been raised to this project. One is that it would not convince Laetrile proponents if no responses or only a few were found. However, the purpose of the retrospective case record review was not to convince Laetrile proponents. It was to determine whether enough evidence of possible antitumor activity can be documented to justify undertaking a prospective clinical trial.

Because of the continued public interest in Laetrile and the suggestion that at least a few objective regressions were associated with the drug in the retrospective study, a prospective evaluation trial of Laetrile will be initiated. Both phase II and III trials will be carried out.

## REFERENCES


1. Laster WR, Jr, and Schabel FM, Jr. Experimental studies of the antitumor activity of amygdalin MF (NSC-15780) alone and in combination with β-glucosidase (NSC-128056). Cancer Chemother Rep 59:951–965, 1975.
2. Wodinsky I, and Swiniarski JK. Antitumor activity of amygdalin MF (NSC-15780) as a single agent and with β-glucosidase (NSC-128056) on a spectrum of transplantable rodent tumors. Cancer Chemother Rep 59:939–950, 1975.
3. Stock CC, Martin DS, Sugiura K, et al. Antitumor tests of amygdalin in spontaneous animal tumor systems. J Surg Oncol 10:89–123, 1978.
4. Stock CL, Tarnowski GS, Schmid FA, et al. Antitumor tests of amygdalin in transplantable animal tumor systems. J Surg Oncol 10:81–88, 1978.
5. Laetrile. Commissioner's Decision on Status. Federal Register, Friday, Aug 5, 1977. Part II. Washington, DC, Department of Health, Education, and Welfare, Food and Drug Administration, pp 39,768–39,806.
6. Smith FP, Butler TP, Cohan S, et al. Letter: Laetrile toxicity: a report of two cases. JAMA 238:136, 1977.
7. Davignon JP. Contaminated Laetrile: a health hazard. N Engl J Med 297:1355–1356, 1977.
8. Lipsett MD, and Fletcher JC. Ethics of Laetrile clinical trials. N Engl J Med 297:1183–1184, 1977.
9. Moertel CG. A trial of Laetrile now. N Engl J Med 298:218–219, 1978.
10. Newell GR. Why the National Cancer Institute chooses a case-review of Laetrile. N Engl J Med 298:216–218, 1978.
11. Ellison NM, Byar DP, and Newell GR. Special report on Laetrile: the NCI Laetrile review. N Engl J Med 299:549–552, 1978.

# Clinical Considerations in the Design of Clinical Trials

**Stephen K. Carter** [1,*]

## SUMMARY

**The clinical trial is the tool used to improve the survival and quality of life for cancer patients through clinical research. The essence of any good clinical trial is an idea worth testing. The multiplicity of questions that can be asked in clinical research requires flexibility in the design of clinical trials. The controversy that rises between randomized trials versus historical control trials at times ignores the reality that different questions require different clinical trial methodologies. Whether a trial is randomized or not should not be viewed as an ethical question, but instead as a strategic question. One way to look at clinical trials is to divide them into exploratory, confirmatory, and explanatory studies. The exploratory study is an initial study which looks at a new idea. The confirmatory study attempts to replicate the results of a positive exploratory study. The explanatory study takes an established regimen and attempts to modify and/or dissect it to improve its therapeutic index. Each of these trial types requires a different mix of strategies which must also be impacted upon by the exigencies of a disease-oriented approach to clinical research. Long-term adjuvant chemotherapy trials are new and long-term analysis must be a sine qua non. While the initial analysis will be relapse-free survival versus acute toxicity, the ultimate analysis will entail balancing overall survival with chronic toxicity.**

[Cancer Treat Rep 64:367–371, 1980]

The object of clinical research in oncology is to improve both the survival and the quality of life in patients with cancer. The methodology that clinical research utilizes to accomplish this aim is the clinical trial. All clinical trials involve three major human elements: the patient, the physician, and the statistician. The patient is the currency of the clinical trial, and at all times the trial should deliver to him or her the best available care while accumulating data which will be of value in the future.

The essence of any good clinical trial is an idea that is worth testing. That idea may be a hypothesis for improvement of therapy, a new regimen, or a new drug. The testing of an idea goes through several phases. If it is a new drug it undergoes phase I, II, and III testing (1). If it is a new regimen it initially undergoes a pilot study which usually encompasses phase I and II testing for new drugs. If the pilot study is encouraging, further (phase III) study is undertaken. All of these various trials take place in patients who have one of the > 100 diseases called cancer. The data that accrue have to be analyzed within the framework of the existing data base for that particular kind of cancer. The design of the trial has to fit into the strategy of the overall attack which is considered optimal for that tumor. Given the heterogeneity of ideas to be tested, the various kinds of sequential testing an idea can flow through, and the myriad types of cancer, it should be obvious that no single clinical trial design approach will be adequate.

One of the major controversies in oncologic clinical research today is between the proponents of the prospectively randomized clinical trial and the proponents of the historically controlled trial. The supporters of historical controls emphasize that all knowledge is historical and that most major gains in the history of medicine have been achieved by trials which compared new findings to what was previously known. These supporters react strongly to what they consider the religious fervor with which some supporters of the randomized control trial attack them and defend the faith. They also assert that

[1]Northern California Cancer Program, Stanford University, and University of San Francisco, CA.

**Reprint requests to:* Stephen K. Carter, MD, Northern California Cancer Program, PO Box 10144, Palo Alto, CA 94304.

the historical control makes clinical trials more efficient since all patients can be exposed to the new idea to be tested and the answer can be achieved with greater dispatch.

The supporters of randomized clinical trials emphasize the manifold variables which can influence the results of a clinical trial: the heterogeneity of clinical material, the selection biases inherent in therapy choices, and the lack of comparability in many sequential studies from groups and institutions using identical regimens. They believe that only randomization can neutralize some of the patient selection biases and achieve a degree of comparability which will lend confidence to findings of superiority for a given type of treatment. Since randomization itself does not ensure comparability, stratification for important prognostic variables is often used in trial design. However, even with stratification and randomization, comparability is not always achieved. For this reason, the proponents of randomized trials assert that the risk of failing to achieve comparability, with a consequent spurious result, is that much greater if historical controls are used. They also point out that since historical controlled studies are rarely accepted alone and are often repeated, in the long run, they represent a less efficient use of patient resources.

At times the debate becomes intense and generates more heat than light. One extreme view is that all studies which are nonrandomized are by definition unscientific. The opposite view is that all randomized trials are antiintellectual. In all probability the truth lies somewhere in between. Given the myriad kinds of ideas to be tested and the manifold clinical possibilities for their testing, there must be situations where a sequential study relating findings to an appropriate historical data base will be optimal, and other situations which will clearly call for prospectively randomized controls. The debate in clinical trials should specify where one or the other or both approaches would be optimal.

The potential ideas to be tested in oncologic clinical trials can come from many sources. The most common sources are the various modalities utilized to treat cancer (*see* fig 1). Each of these ideas can have a nearly infinite number of subsets. Chemotherapy is the most striking example (2). If one assumes approximately 30 standard drugs and 50 investigational drugs, each of which has various dose schedules, levels, and routes, one could devise two-, three-, four-, and five-drug combinations which could easily run into the billions. With radiation therapy there are variables of dose, fractionation, delivery techniques, and port, all of which could be combined with sensitizing drugs or any of a large number of potential chemotherapeutic regimens. Surgery and immunotherapy have their manifold variables and so make their own contribution in an almost infinite number of possibilities for combined modality therapy.

1. New surgical techniques
2. New radiation delivery techniques
3. New radiation-fractionation schedules
4. Radiation sensitizers
5. Hyperthermia
6. New drug classes
7. New analogs of existing known actives
8. New combination chemotherapy regimens
9. New drug schedules
10. New immunotherapeutic approaches
11. Surgery plus radiation
12. Surgery plus chemotherapy
13. Surgery plus immunotherapy
14. Radiation plus chemotherapy
15. Radiation plus immunotherapy
16. Chemotherapy plus immunotherapy
17. Biologic markers and their impact
18. New staging techniques
19. New followup techniques
20. Toxicity modifiers

FIGURE 1.—20 possible general modality-oriented ideas to be tested in current clinical research.

All of these possibilities have their final common pathway in the therapy of one of the many diseases called cancer. While the initial trials (phase I or pilot) may be modality-oriented, the ultimate trials (phase II and III) will be as much disease-oriented, if not more so. In view of the wealth of modality-oriented possibilities, the choice for study should be made by a multidisciplinary team of investigators. It should also be obvious that with this abundance of possibilities no one clinical trial design will be optimal for all.

Another way to look at clinical trials is to divide them into exploratory, confirmatory, and explanatory studies (table 1). The exploratory study is an initial study which looks at a new idea. If a totally new drug is being tested, the exploratory study is a phase I study. If a new combination regimen is being tested, it could be called a pilot study or a phase I and II study. This kind of study will utilize a relatively small number of patients and will tend to be modality-oriented. In most cases it will not require a prospectively randomized clinical trial design. The endpoint of such a trial will be a decision to continue or abandon the study. The results of the study will be compared to the existing massive historical data base. If the new idea appears to offer a reasonable possibility of improving the results that have been reflected in the historical data base, further studies will be called for. Positive results from an exploratory study may render unethical a prospectively randomized clinical trial comparing the new idea to the best current standard treatment. If a new drug regimen for extensive colorectal carcinoma has given 15 consecutive complete remissions, most investigators would recommend an uncontrolled sequential confirmatory study. However, if the same reg-

TABLE 1.—3 conceptual clinical trials

| Orientation | | | | Randomization | |
|---|---|---|---|---|---|
| Modality | Disease | | | Yes | No |
| + + + | + | Exploratory study | {Pilot study or phase I–II study} | ± | + + + |
| + | + + + | Confirmatory study | {Phase II–III study} | + + + | + |
| + | + + + | Explanatory study | {Phase III or IV study} | + + + | ± |

imen has given three, five, or seven responses, many of which were partial, the debate would be more heated. With these lower response rates it surely would not be an ethical problem *not* to give the new regimen to the next several patients seen. On the other hand, standard treatment of this disease with fluorinated pyrimidines gives results which generally are so poor that it would be quite ethical to withhold this therapy (3). Therefore, the testing of the new regimen in a prospectively controlled comparison with fluorinated pyrimidines is in truth not an ethical question but a strategic one. It is unfortunate that the proponents of various positions in the clinical trials debate tend to use the emotionally charged rhetoric of ethics as they defend their own position and attack their opponents.

This problem of interpreting the positivity of an exploratory study with regard to the clinical trial design strategy to be used as followup is more crucial in the combined modality arena. Combined modality exploratory studies often cannot be carried out in the therapeutic setting in which they are ultimately expected to be of value. The only "pilot" setting for drugs to be used in surgical adjuvant trials has been in the patient with advanced disease of the same tumor type or in the mouse. It has been assumed that drugs which work in advanced disease will be more effective in a surgical adjuvant setting where residual tumor cell volume is believed to be lower and cell kinetics is possibly more favorable (4). These trials in advanced disease define the drugs to be used and indicate feasible schedules and dose levels. The rodent tumor systems provide a comforting experimental rationale. Again, it should be obvious that no result achieved with these kinds of exploratory studies can ethically preclude a prospectively controlled study. In fact, the possible harm from adjuvant drug therapy and the potential socioeconomic and psychosocial impact of recommending years of drug therapy routinely after surgery should persuade us to insist on a prospectively controlled study.

The confirmatory study is the trial which attempts to replicate the results of a positive exploratory study and so establish its applicability to a wider trial or to direct clinical application. In chemotherapy these trials tend to fall into the phase II and III range. One hallmark of science is reproducibility and therefore, when a new regimen is reported to be positive it must be confirmed, not only by the team that reported it initially but by several other groups as well. If such a study confirms a highly active new regimen for advanced disease, a sequential study may be indicated. The MOPP (mechlorethamine, vincristine, prednisone, and procarbazine) regimen for stage III and IV Hodgkin's disease was initially tested in 43 consecutive patients without a concomitant control (5). Its confirmation by the major cooperative groups studying Hodgkin's disease in the US took a variety of forms (6). One group treated several hundred patients with MOPP without any concomitant control (viz patients receiving what had been standard therapy) and confirmed the high complete response rate. A second group compared MOPP to mechlorethamine alone in a prospectively controlled design, confirmed MOPP's high activity, and showed that its superiority was statistically significant. A third group compared MOPP to another multidrug regimen in a prospectively randomized study; in this study the investigators sought confirmation of MOPP's activity as well as fresh data on the new regimen, which had mainly small pilot exploratory data to support it. Once again the MOPP regimen was shown to be superior. All three approaches confirmed the initial MOPP exploratory study, and for these groups MOPP now became the standard best regimen from which to launch new attempts at improved control of Hodgkin's disease. All of these clinical trial designs were valid and none were "unethical."

One of the unresolved questions is how many confirmatory studies are needed before a regimen becomes established; no set number, however, will be applicable to every positive idea that is generated. In oncology, confirmatory studies seem to be more numerous than the principles of design would justify. Furthermore, many studies provide only imperfect confirmation because each group has felt impelled to tinker with the regimen somewhat. The mass of contradictory results so produced is difficult to interpret since the regimen being tested is in truth a variety of noncomparable regimens.

The explanatory study is one which takes an established regimen and attempts to modify and/or dissect it to improve its therapeutic index. While some of these modifications must pass through an exploratory phase of dose finding, this is usually brief. In chemotherapy the explanatory studies are usually of the phase III variety, and since they usually are not designed to demonstrate large differences of effect, they mainly involve large numbers of patients in a prospectively randomized design.

One example of such a study was one by the Cancer and Acute Leukemia Group B on Hodgkin's disease which compared MOPP to the same regimen without mechlorethamine, the same regimen with BCNU substituted for mechlorethamine, and the three-drug regimen of BCNU, vincristine, and prednisone (6). Other examples are the numerous breast cancer (advanced disease) studies which dissected the five-drug regimen originally described by Cooper, testing three- and four-drug regimens such as CMF (cyclophosphamide, methotrexate, and 5-fluorouracil), CFP (cyclophosphamide, 5-fluorouracil, and prednisone), and CMF plus prednisone (7).

We have effective treatments for cancer and have a wide range of potential ways of putting these effective treatments together which we hope will make them more effective. In my opinion, we require properly designed controlled clinical trials to delineate levels of superiority with various therapeutic approaches, so that we can establish a firm footing for the next therapeutic assault on cancer. If generally accepted by practicing clinicians, many of the new treatments now being evaluated would constitute a major change in cancer care and would greatly increase its expense and risk to the patient; one example is adjuvant therapy after surgery and/or radiation therapy for local and regional tumor control. Before recommending 2 years of systemic treatment with drugs or immune-modulating vaccines for every patient with breast or colorectal cancer, we should be sure that the expense (both economic and psychologic) and risk have been justified by clinical trials which clearly establish the value of this new approach. If this approach will only increase cures by 20%–30% (an increase which would be highly significant), we need the discipline and the scientific security of the controlled clinical trial.

Long-term analysis is a sine qua non in clinical trials of adjuvant chemotherapy. In these trials the participating chemotherapist (who in many cases is the study chairman) does not have objective regression or response to aid his assessment. What he is treating are microscopic foci of metastatic disease, which he assumes from theoretic assumptions to be present, but which he cannot see, measure, or evaluate except indirectly in terms of ultimate relapse and survival. In this situation, the first endpoint in adjuvant drug studies is the relapse rate, and from this can be constructed an actuarial projection of a relapse-free survival curve. The first available cost-benefit analysis is between relapse rate and acute toxicity. What is not known at this time is the minimal duration of followup which would ensure that an actuarial projection of a relapse-free survival, based on relapse rates at early fixed points in time, will predict the ultimate outcome. Since new trials must be designed quickly as case accession to studies approaches the numbers required, the need and pressure for an early analysis become severe. Over the last few years the preliminary literature has contained reports of adjuvant trials in breast cancer (8,9) and osteogenic sarcoma (10,11) which have caused a great deal of excitement, attracted much publicity, and generated new adjuvant trials. The authors of these papers stressed that their results were preliminary, but unfortunately their conservatism was not matched by those who seized upon these data as if they firmly established the value of adjuvant chemotherapy and supported its theoretic assumptions. In any event, the enthusiasm so generated led clinical investigators to develop new protocols and encouraged practicing clinicians to accept adjuvant chemotherapy as an established practice. Subsequent analysis of these studies has shown that the cautionary note sounded by their authors was fully justified (12–15). With respect to the early positive actuarial predictions, late relapses have lowered the overall differences between treated and control groups, and it is sobering to contemplate that in terms of overall survival we have a long way to go before we can make a confident statement about the ultimate value of these regimens. The principal investigators of the newer adjuvant trials will probably not publish their preliminary results as early as the investigators in the original breast cancer and osteogenicsarcoma protocols. If they do, we hope that the editorial writers and general readers will be more cautious this time.

When faced with a devastating and often fatal disease such as cancer, there is a great thrust to find promising new treatments and a great willingness to accept them. There is always the hope of finding a breakthrough therapy which will have a dramatic impact on mortality. Should such a breakthrough be found, its clinical evaluation will be simple and will not require any complex clinical trial designs or extensive sophisticated analytic techniques. Many advances in medical practice have occurred and will continue to occur in small steps. There is a need for professionals, politicians, and the public to understand that progress is often hard to discern and difficult to separate from influences other than innovation. Yet, great progress results from the aggregation of these limited advances over time.

## REFERENCES


1. Carter SK. Clinical trials in cancer chemotherapy. Cancer 40:544–557, 1977.
2. Carter SK, Bakowski M, and Hellman K. Chemotherapy of Cancer. New York, NY, John Wiley, 1977.

3. CARTER SK. Guest Editorial: Large bowel cancer—the current status of treatment. JNCI 56:3–10, 1976.
4. CARTER SK, and SOPER WT. Integration of chemotherapy into combined modality treatment of solid tumors. Cancer Treat Rev 1:1–13, 1974.
5. DEVITA VT, JR, SERPICK AA, and CARBONE PP. Combination chemotherapy in the treatment of advanced Hodgkin's disease. Ann Intern Med 73:881–895, 1970.
6. GOLDSMITH MA, and CARTER SK. Combination chemotherapy of advanced Hodgkin's disease—a review. Cancer 33:1–8, 1974.
7. CARTER SK. Integration of chemotherapy into combined modality treatment of solid tumors. VII. Adenocarcinoma of the breast. Cancer Treat Rev 3:141–174, 1976.
8. BONADONNA G, BRUSAMOLINO E, VALAGUSSA P, ET AL. Combination chemotherapy as an adjuvant treatment in operable breast cancer. N Engl J Med 294:405–410, 1976.
9. FISHER B, CARBONE P, ECONOMOU SG, ET AL. L-Phenylalanine mustard (L-PAM) in the management of primary breast cancer. A report of early findings. N Engl J Med 292:117–122, 1973.
10. CORTES EP, HOLLAND JF, WANG JJ, ET AL. Amputation and Adriamycin in primary osteosarcoma. N Engl J Med 291:998–1000, 1974.
11. JAFFE N, FREI E III, TRAGGIS D, ET AL. Adjuvant methotrexate and citrovorum-factor treatment of osteogenic sarcoma. N Engl J Med 291:994–997, 1974.
12. BONADONNA G, ROSSI A, VALAGUSSA P, ET AL. Adjuvant chemotherapy with CMF in breast cancer with positive axillary nodes. *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 83–95.
13. CORTES EP, HOLLAND JF, and GLIDEWELL O. Effects of Adriamycin in the adjuvant treatment of osteosarcoma. *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 407–423.
14. FISHER B, and REDMOND C. Studies of the National Surgical Adjuvant Breast Project (NSABP). *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 67–81.
15. FREI E III, JAFFE N, SKIPPER HE, ET AL. Adjuvant chemotherapy of osteogenic sarcoma: progress and perspectives. *In* Adjuvant Therapy of Cancer (Salmon SE, and Jones SE, eds). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1977, pp 49–64.

# Use of Prognostic Factors in Analysis of Historical Control Studies [1]

Edmund A. Gehan,* Terry L. Smith, and Aman U. Buzdar [2]

## SUMMARY

**This paper considers the use of regression methods for the interpretation of comparative results between treatments in historical control studies. A logistic regression model is used for response rate studies and Cox's regression model for studies of disease-free or overall survival. For both models, stepwise regression procedures may be utilized to test for treatment differences adjusting for differences in patient characteristics between the treatment groups. Examples are given from two clinical studies: the first example concerned a study which compared disease-free and overall survival experience between patients receiving surgery plus radiation therapy plus FAC-BCG (5-fluorouracil, Adriamycin, and cyclophosphamide-bacillus Calmette-Guérin) as adjuvant treatment compared with surgery plus radiation therapy alone for patients with stage II or III breast cancer; the second example concerned a study of acute leukemia in which differences among complete remission rates by institution were interpreted using a logistic regression equation to adjust for prognostic factors.**

[Cancer Treat Rep 64:373–379, 1980]

The purpose of this paper is to outline some of the statistical methodology for utilizing knowledge of prognostic factors in the analysis of clinical studies and to give examples of its successful applications. Examples of applications will be from studies of breast cancer (1) and acute leukemia.

This paper was part of a session on "Basic Design Considerations" in clinical studies. The simplest design is a historical control study in which all patients in a consecutive series receive a new treatment and comparison is made with results for a control treatment in the recent past. Administering the new treatment to a series of patients permits the clinical trial to proceed more rapidly than randomizing patients to two or more treatments, eliminates ethical problems because all patients are offered the proposed new treatment, and makes possible better estimates of the effectiveness of the new treatment in a given period of time. Utilizing objective endpoints in the analysis and confirming good results obtained in additional patients and by other investigators provide a convincing mechanism for evaluating the effectiveness of new treatments. The only real difficulty with historical control studies lies in analyzing comparative results between the new and standard treatment and making appropriate adjustments for differences in the prognostic characteristics of the patients between groups.

It is assumed in this paper that a consecutive series of patients has received a new treatment, and comparison will be made with the results obtained from a recent series of patients who have received a control treatment. It is desired that the analysis take into account possible differences in the prognostic factors of the patients receiving each treatment. In a historical control study, the analysis, rather than the design, provides the main focus for taking into account the prognostic characteristics of the patients. Gehan (2) has given reasons for considering historical control studies and has outlined procedures for adjusting for prognostic characteristics (3).

## METHODS

Statistical regression models can be used to relate the probability of complete remission or survival time to patient characteristics. The probability of response may be

[1]Supported by Public Health Service grants CA-11430 and CA-12014 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston.

**Reprint requests to:* Edmund A. Gehan, PhD, Department of Biomathematics, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, 6723 Bertner Ave, Houston, TX 77030.

related to prognostic factors using a logistic regression model. The form of this model is:

$$\log_e \left( \frac{p_i}{1-p_i} \right) = a_0 + a_1 (x_{1i} - \bar{x}_1) + \ldots + a_k (x_{ki} - \bar{x}_k),$$

where $p_i$ is the probability of response for the $i^{th}$ patient, the a's are regression coefficients to be estimated, the $x_i$'s are patient characteristics possibly related to response, and the $\bar{x}$'s are the mean values of each patient characteristic. This model can be fitted in stepwise fashion so the first variable included is that giving the best prediction of probability of response; the second variable, when added to the first, gives the best pair of variables for predicting response, and so on. The theory for fitting this model to experimental data is given by Cox (4). A computer program for fitting the model in a stepwise manner has been given by Lee (5). For relating survival time or disease-free interval to patient characteristics, Cox's regression model has been used. The form of this model is:

$$\log_e \left( \frac{\lambda_i(t)}{\lambda_0(t)} \right) = b_1 (x_{1i} - \bar{x}_1) + \ldots + b_k (x_{ki} - \bar{x}_k),$$

where $\lambda_i(t)$ is the hazard function (or roughly the risk of failure per unit time) at time t for the $i^{th}$ individual, the b's are regression coefficients fitted in the model, the $x_i$'s are patient characteristics, and the $\bar{x}$'s are mean values. This model can also be fitted in forward stepwise fashion so the only x variables included are those contributing to the prediction of survival time or disease-free interval. The theory for using this model has been given by Cox (6). Kay (7) has reviewed recent developments concerning Cox's regression model including methods for testing goodness-of-fit, and has given an example.

In a historical control study in which treatment A was first given in the sequence followed by treatment B, the data on endpoints (response rates or survival experience) will nearly always be more complete on treatment A than treatment B. This is especially true in survival studies when a larger proportion of patients may have died while receiving treatment A compared with treatment B because of longer followup experience. First, in this circumstance, a regression model (either logistic or Cox's regression model depending upon whether the endpoint is probability of response or survival) can be fitted to the group of patients receiving treatment A to determine patient characteristics related to prognosis. Second, another regression model can be fitted to the combined set of patients receiving treatment A and treatment B with the inclusion of another regression variable representing treatment effect (eg, $x_{k+1}$ equals 0 for patients receiving treatment A or 1 for patients receiving treatment B). A chi-square test of significance for the regression coefficient representing treatment effect can be made to determine whether the type of treatment administered adds to the predictability of response or survival. An assumption is needed that the same patient characteristics related to prognosis for treatment A will also be related to prognosis for treatment B. As the data for patients receiving treatment B become more complete, this assumption can be checked. One way would be to fit a model with the same patient characteristics separately for the patients receiving each treatment and to compare regression coefficients between models. If the same patient characteristics are important and the associated regression coefficients are similar, then the utility of this regression approach would be confirmed. Of course, a further assumption is required in a historical control study. It is that possible differences between treatment groups in types of patients, diagnostic procedures, and supportive therapies, are either not related to prognosis or not sufficient to explain differences in therapeutic outcome between groups. An examination should be made of whether any differences exist between treatment groups that may contribute to differences in prognosis.

A general approach in nonrandomized historical control studies holds that the study should be analyzed in several different ways. If all ways lead to the same conclusion, this should strengthen confidence that the conclusion is correct. There are two other ways of using regression models to interpret data from historical control studies. Though interrelated, each of these ways provides a different perspective on the data. When a regression model has been fitted to data from a historical control series, the model may then be used to determine an "index of prognosis." This is done by applying the regression model to patients receiving treatment A or B to obtain a predicted chance of response (from a logistic regression model) or a hazard ratio (from Cox's regression model) for each patient. The hazard ratio is the ratio of the risk of failure per unit time for the $i^{th}$ patient ($\lambda_i(t)$) to the standardized risk for a patient with average values of all patient characteristics ($\lambda_0(t)$). The indices of prognosis can then be grouped into subcategories, and the patients receiving treatment B can be compared to those receiving treatment A within subcategories of patients that have comparable indices of prognosis. Standard chi-square tests may be used to compare response rates or methods for comparing survival distributions (8,9) may be used.

Another method of utilizing regression models determines important patient characteristics related to prognosis and compares differences between treatments within subgroups of patients that have comparable values for the characteristics related to prognosis. For example, if menopausal status and number of nodes are related to disease-free survival for patients with breast cancer, patients receiving the two treatments could be compared

within menopausal status by nodal status subgroups. If it is established that the groups of patients are comparable with respect to prognostic features, then overall comparisons between groups can be made in response rate, disease-free survival, and survival. Note that we are suggesting that the analysis many clinicians often carry out first should be done last.

## EXAMPLES OF REGRESSION METHODS IN CLINICAL STUDIES

This section gives two examples, both of which involve the use of regression models derived from historical control series of data. The first was reported by Buzdar et al (1), and concerns the disease-free and overall survival experience of women with stage II or III breast cancer. The second is concerned with the analysis of differences in results among institutions in the response rates of patients with acute leukemia.

### Breast Cancer Study

This study of adjuvant treatment for breast cancer patients was conducted at The University of Texas System Cancer Center between 1972 and 1976. All patients had stage II or III breast cancer with one or more histologically positive axillary lymph nodes and had definitive surgery usually followed by radiation. Patients in the new treatment group received combination chemotherapy consisting of 5-fluorouracil, Adriamycin, cyclophosphamide, and BCG immunotherapy (designated FAC-BCG treatment). Between January 1974 and October 1976, 131 patients received FAC-BCG as adjuvant treatment. The control patients were drawn from a group of women with breast cancer who were admitted to The University of Texas System Cancer Center between January 1972 and December 1973. This control group included 151 patients who had stage II or III breast cancer and who had at least one involved node at the time of surgery (and for whom the exact number of involved nodes was known). The major endpoints for evaluation in the analysis were disease-free survival time and overall survival following surgery. The complete details of this study are given in Buzdar et al (1).

Characteristics (and codes) available for both groups of patients were: age (years), menopausal status (1 = premenopausal, 0 = other), size of primary tumor (2 = < 3 cm, 4 = 3–5 cm, 7 = > 5 cm), stage of disease (2 = II, 3 = III), location of surgery (1 = M. D. Anderson Hospital, 0 = other), number of involved nodes (2 = < 4, 7 = 4–10, 12 = > 10), race (1 = white, 2 = other), and type of surgery (1 = radical or modified-radical surgery, 0 = other). There was some evidence that prognostic characteristics were not comparable between groups. Characteristics tending to favor the control group were age and menopausal status, since there was a higher percentage of patients who were > 50 years old (68% vs 52%) and peri- or post-menopausal (75% vs 58%). Patients in the FAC-BCG group were favored with respect to size of tumor, stage of disease, and type of surgery. There was a lower percentage of patients in the FAC-BCG group with primary tumors > 5 cm (15% vs 26%) and stage III disease (25% vs 40%) and a higher percentage receiving radical or modified-radical mastectomy (85% vs 55%). The groups were comparable in the distribution of the number of nodes involved.

Cox's regression model was fitted to the disease-free survival data from the 151 control patients in forward stepwise fashion and the following regression model was obtained:

$$\log_e\left(\frac{\lambda(t)}{\lambda_0(t)}\right) = 0.1110\,(\text{number of nodes}-6.2) + 0.8122\,(\text{stage}-2.4) + 0.8720\,(\text{menopausal status}-0.3),$$

The significance levels of the patient characteristics entering the model were: number of nodes ($P < 0.01$), stage of disease ($P = 0.02$), and menopausal status ($P < 0.01$). Favorable characteristics were: small number of nodes, stage II disease, and postmenopausal status.

Cox's regression model was then fitted to the combined data from FAC-BCG and control patients to test whether adjuvant treatment affected disease-free survival. Type of treatment (0 = control, 1 = FAC-BCG) was included as a variable in the model, in addition to number of involved nodes, stage of disease, and menopausal status. The equation obtained by fitting the model to the combined group of control and FAC-BCG patients was:

$$\log_e\left(\frac{\lambda(t)}{\lambda_0(t)}\right) = -1.6962\,(\text{treatment}-0.5) + 0.0965\,(\text{number of nodes}-6.3) + 0.8616\,(\text{menopausal status}-0.3) + 0.6563\,(\text{stage}-2.3).$$

The first variable entering the model in the forward stepwise fitting process was type of treatment ($P < 0.01$), followed by number of involved nodes ($P = 0.01$), menopausal status ($P = 0.01$), and stage of disease ($P = 0.01$). Type of treatment was the variable most significantly related to disease-free survival, with patients in the control group having 5.5 times the risk of relapse per unit time as patients receiving FAC-BCG.

The regression model fitted to data from the control group only (ie, with the variables, number of nodes, stage of disease, and menopausal status) was applied to the data from both the control and FAC-BCG groups to obtain the hazard ratio, $\lambda(t)/\lambda_0(t)$, in each group of patients. The median hazard ratios were 0.875 in both control and FAC-BCG groups, suggesting that differences in the distribution of prognostic factors between groups tended to balance out so that the overall risk of relapse per unit time was comparable. Hence, an overall comparison in disease-free survival was made between groups. As of

TABLE 1.—Data from Cox's regression model fit to disease-free survival for control patients

| No. of variables in equation | Variables | Value of maximized log-likelihood | Chi-square ($X^2$) value* | Variables giving significantly poorer ($P < 0.05$) fit to disease-free survival than best-fitting variable |
|---|---|---|---|---|
| 1 | No. of nodes | −317.859 | — | |
| | Disease stage | −318.437 | 1.156 | |
| | Premenopausal status | −319.272 | 2.826 | |
| | Tumor size | −319.394 | 3.070 | |
| | Surgery | −319.987 | 4.256 | Surgery |
| | Age | −320.120 | 4.522 | Age |
| | Race | −320.512 | 5.306 | Race |
| | Location of surgery | −320.779 | 5.840 | Location of surgery |
| 2 | No. of nodes + disease stage | −315.135 | — | |
| | No. of nodes + premenopausal status | −315.872 | 1.474 | |
| | No. of nodes + surgery | −316.007 | 1.744 | |
| | Disease stage + premenopausal status | −316.246 | 2.222 | |
| | No. of nodes + tumor size | −316.508 | 2.746 | |
| | Premenopausal status + tumor size | −317.286 | 4.302 | Premenopausal status + tumor size |
| | Premenopausal status + surgery | −317.761 | 5.252 | Premenopausal status + surgery |
| | Disease stage + surgery | −318.260 | 6.250 | Disease stage + surgery |
| | Disease stage + tumor size | −318.408 | 6.546 | Disease stage + tumor size |
| 3 | No. of nodes + disease stage + premenopausal status | −312.254 | — | |
| | No. of nodes + premenopausal status + surgery | −312.759 | 1.070 | |
| | No. of nodes + premenopausal status + tumor size | −313.844 | 3.180 | |
| | No. of nodes + disease stage + surgery | −314.507 | 4.506 | No. of nodes + disease stage + surgery |
| | No. of nodes + disease stage + tumor size | −315.135 | 5.762 | No. of nodes + disease stage + tumor size |
| | No. of nodes + tumor size + surgery | −315.480 | 6.452 | No. of nodes + tumor size + surgery |
| | Disease stage + premenopausal status + surgery | −315.651 | 6.794 | Disease stage + premenopausal status + surgery |
| | Disease stage + premenopausal status + tumor size | −316.148 | 7.788 | Disease stage + premenopausal status + tumor size |
| | Premenopausal status + tumor size + surgery | −316.466 | 8.424 | Premenopausal status + tumor size + surgery |
| | Disease stage + tumor size + surgery | −318.247 | 11.986 | Disease stage + tumor size + surgery |
| 8 | | −310.648 | | |

*Value of $X^2$ for test of difference in maximized log-likelihoods between particular variable (eg, stage) and variable(s) with the largest maximized log-likelihood (eg, nodes).

April 1978 the estimated percentages of patients disease-free at 2 years were 85% in the FAC-BCG group and 64% in the control group; there was a highly statistically significant difference between the disease-free survival curves ($P < 0.01$). Comparisons of the FAC-BCG and the control groups were also made within subgroups of patients comparable in hazard ratio and individual patient characteristics, and these analyses confirmed the highly statistically significant overall difference between FAC-BCG and control patients.

A question arose concerning the dependency of the results on the forward stepwise fitting process of the regression model and whether other variables could be found that were equally predictive of disease-free survival. Table 1 gives values of maximized log-likelihood as determined with one, two, or three patient characteristics in the regression equation. For a fixed number of variables in the regression equation, minus two times the difference in maximized log-likelihood is distributed approximately as chi-square with one degree of freedom. Comparison of this statistic with 3.84 indicates whether there are statistically significant (ie, $P \leq 0.05$) differences in the fit of the regression models to the data. The last column in table 1 gives the chi-square value corresponding to a test of the difference in regression fit between a model with a particular variable (or set of variables) versus the best-fitting single variable (or sets of two or three variables). For example, with one variable in the regression equation, the largest maximized log-likelihood is associated with number of nodes, indicating that a single variable is the best predictor of disease-free survival. The chi-square value of 2.826 for premenopausal status indicates that that variable does not give a significantly poorer fit than one based upon number of nodes, while an equation with only a single term for race would give a significantly poorer fit than one based upon number of nodes (since chi-square equals 5.306 for race).

Table 1 shows that the best individual predictors of

TABLE 2.—Cox's regression model fit to combined control and FAC-BCG groups relating disease-free survival to patient characteristics with and without treatment

| Variables in regression equation | Maximized log-likelihood | Test of significance for FAC-BCG vs control |
|---|---|---|
| No. of nodes, menopausal status, disease stage | −470.933 | $P < 0.01$ |
| No. of nodes, menopausal status, disease stage, treatment | −456.070 | |
| No. of nodes, menopausal status, surgery | −473.529 | $P < 0.01$ |
| No. of nodes, menopausal status, surgery, treatment | −460.528 | |
| No. of nodes, menopausal status, tumor size | −475.591 | $P < 0.01$ |
| No. of nodes, menopausal status, tumor size, treatment | −460.034 | |

disease-free survival are: number of nodes, stage of disease, menopausal status, and size of tumor. When sets of two variables are considered, the combinations of these variables and number of nodes plus type of surgery give the best pairs of variables. Table 1 shows that the best fitting sets of three variables include number of nodes and menopausal status, together with stage of disease, type of surgery, or size of tumor.

To consider whether the significance of treatment effect depended upon the particular set of three patient variables chosen, type of treatment was added as a potential fourth variable to each of the best-fitting sets of three variables. The data given in table 2 indicate that type of treatment adds significantly to the prediction of disease-free survival, no matter which of the three best-fitting sets of three variables are chosen. Hence, this analysis shows that FAC-BCG was better than control treatment when adjustments were made for differences between groups in number of nodes, premenopausal status, stage of disease, type of surgery, and size of tumor.

Cox's regression model was also fitted to the survival data from the control group and the patient characteristics most related to survival were menopausal status, size of primary tumor, and number of nodes. When the regression model was fitted to the combined group of data from FAC-BCG and control patients, it was demonstrated that type of treatment was the first characteristic to enter the model ($P < 0.01$) and the risk of death per unit time for a patient receiving no adjuvant chemotherapy treatment was 6.6 times that for a patient receiving adjuvant treatment with FAC-BCG. As of April 1978, the estimated percentages of patients surviving at 2 years were 97% in the FAC-BCG group and 84% in the control group; the overall difference between survival curves was highly statistically significant ($P < 0.01$).

This was a historical control study in which the groups of patients were not comparable with respect to certain important characteristics. However, use of Cox's regression model enabled a test to be made of the difference between treatments that was adjusted for the differences in patient characteristics. The consistency in the results for disease-free survival and overall survival and the analysis of patients in various subgroups strongly suggest that the overall prognosis is significantly superior for patients receiving FAC-BCG.

The major questions usually asked of historical control studies concern the selection of patients for the control group. That is, was there bias in choosing unfavorable patients for the control group or was there some evidence that prognosis for patients with a given disease was improving with chronologic time. Table 3 compares the control patients in this study and the control patients from

TABLE 3.—% of patients disease-free at 2 yrs in 3 control groups

| | Control group | | |
|---|---|---|---|
| | L-PAM | CMF | FAC-BCG |
| Patient characteristic | No. of patients (%) | No. of patients (%) | No. of patients (%) |
| Age < 50 yrs | 60 (59) | 86 (60) | 38 (56) |
| No. of nodes | | | |
| 1–3 | 31 (71) | — (70) | 14 (71) |
| ≥ 4 | 29 (49) | — (12) | 24 (42) |
| Age ≥ 50 yrs | 109 (74) | 93 (75) | 113 (74) |
| No. of nodes | | | |
| 1–3 | 56 (86) | — (80) | 46 (80) |
| ≥ 4 | 53 (62) | — (60) | 67 (69) |
| Totals | 169 (68) | 179 (64) | 151 (69) |

two other randomized studies using adjuvant treatment for breast cancer, one by Fisher et al in Pittsburgh (10) using L-phenylalanine mustard (L-PAM), and the other by Bonadonna et al in Milan (Italy) (11) using cyclophosphamide, methotrexate, and 5-fluorouracil (CMF) as adjuvant treatment. When the control groups are examined for the distributions of age and number of nodes, the control group in this study compares well with the control groups used in the other studies. The only subgroup in which there was an exception was in the Bonadonna study for patients < 50 years old who had more than four nodes. The estimated percentage of patients disease-free at 2 years was only 12%; however, the number of patients involved was small, and overall there was much more evidence of consistency than inconsistency among control groups.

To determine whether there was evidence that prognosis was improving with chronologic time, the control group of patients was divided into two subgroups, those entering before 1973 and those entering in 1973 or later. For the first 12 months the disease-free survival curves are comparable, but a separation occurs thereafter, so that the percentage of patients disease-free at 2 years was 71% in the early control group compared with 56% in the later control group. Overall, there was a statistically significant difference between disease-free survival curves favoring the early control group. Hence, these data do not provide evidence that prognosis was improving with time; rather, there is some suggestion that it was worsening. The difference between curves cannot be explained precisely in terms of prognostic characteristics, since the indices of prognosis based upon hazard function were comparable in the two groups. Possible reasons for differences include better methods for detecting recurrences and more determined followup of patients in the more recent control group. In any event, the FAC-BCG group results were significantly better than those in either the early or the later control groups, but these results indicate the precautions that should be taken in analyzing and interpreting historical control studies.

## Acute Leukemia Study

This analysis was concerned with the interpretation of differences among response rates at institutions participating in a chemoimmunotherapy study of acute leukemia (designated the CIAL study) in the Southwest Oncology Group (SWOG No. 7416/17). A logistic regression equation was used to interpret differences in response rates among institutions, taking into account the prognostic characteristics of the patients. The logistic regression equation was derived from previous studies of continuous infusion with 5- or 10-day OAP (vincristine, cytosine arabinoside, and prednisone). The logistic regression equation was:

$$\log_e\left(\frac{p_i}{1-p_i}\right) = 0.02686 - 0.04236\ (\text{age}-44.031) - 0.59297\ (\text{infection status}-0.370) - 0.35854\ (\text{AML}-0.872) - 0.01431\ (\text{hgb}-9.155) - 0.02018\ (\log\ (\text{wbc})-4.127),$$

where $p_i$ is the predicted complete remission rate for a patient based on the five prognostic characteristics. The codes for the variables were: age (years), infection status (0 = no, 1 = yes), acute myelocytic leukemia (AML) (0 = no, 1 = yes), hemoglobin value, and log wbc. For patients in the CIAL study, information available at the start of the study concerning age, infection status, type of leukemia, hemoglobin value, and logarithm of the wbc count was used to obtain a predicted probability of response based upon the previous 5- and 10-day OAP data.

Table 4 gives observed and expected response rates from 16 institutions entering ten or more patients into the CIAL study. The overall complete remission rate was 57% in 419 patients. The observed complete remission rate was 67% at institution A, the only one that was ≥ 10% above the average complete remission rate of 57%. Institution M had the lowest complete remission rate of 21% and institutions B, C, H, and M were ≥ 10% below the average response rate. Comparisons among institutions differed markedly when observed minus average complete remission rates were calculated compared with observed minus predicted complete remission rates based upon the 5- and 10-day OAP data. Overall, the standard deviation of observed minus average response rates was 12%, compared with 9% for observed minus predicted complete remission rates. For observed minus predicted complete response rates, institution G had a response rate of 64% in 89 patients; this was 19% above their predicted response rate of 45%. This institution had a higher response rate than average, while their predicted response rate was substantially less than average. Institution K, with a response rate of 61%, was also > 10% above predicted. Institutions H and M were ≥ 10% below predictions. The high response rate at institution A can be best explained by the more favorable patients entering their study.

If one considers the size of the institutions, four (E, G, K, and O) entered ≥ 30 patients. These institutions averaged 7% above their predicted complete remission rates and included the only two institutions that were ≥ 10% above predicted. The remaining 12 institutions had complete remission rates averaging 1% under predicted values.

Based upon these relatively small samples of patients, these data suggest that observed complete remission rates do differ in an important way by institution, but much less so when observed minus predicted response rates are considered. The size of an institution may have some influence on response rate, but this point requires further investigation.

TABLE 4.—Observed vs predicted response rates for CIAL study by institution

| Institution | Total No. of patients | No. of patients with complete remission (%) | | % observed minus predicted | % observed minus average |
|---|---|---|---|---|---|
| | | Observed | Predicted | | |
| A | 12 | 8 (67) | 7.1 (60) | +7 | +10 |
| B | 19 | 9 (47) | 8.6 (46) | +1 | –10 |
| C | 13 | 6 (46) | 5.8 (46) | 0 | –11 |
| D | 19 | 11 (58) | 10.8 (57) | +1 | +1 |
| E | 32 | 16 (50) | 16.6 (52) | –2 | –7 |
| F | 29 | 18 (62) | 16.6 (58) | +4 | +5 |
| G | 89 | 57 (64) | 39.4 (45) | +19 | +7 |
| H | 21 | 8 (38) | 10.1 (49) | –11 | –19 |
| I | 27 | 17 (63) | 17.5 (65) | –2 | +6 |
| J | 29 | 15 (520 | 13.4 (47) | +5 | –5 |
| K | 36 | 22 (61) | 17.9 (50) | +11 | +4 |
| L | 13 | 8 (62) | 8.7 (67) | –5 | +5 |
| M | 14 | 3 (21) | 5.6 (41) | –20 | –36 |
| N | 14 | 9 (64) | 7.6 (55) | +9 | +7 |
| O | 36 | 22 (61) | 21.9 (61) | 0 | +4 |
| P | 16 | 8 (50) | 8.5 (53) | –3 | –7 |
| Totals | 419 | 237 (57) | 216.1 (52) | | |

## REFERENCES


1. BUZDAR A, GUTTERMAN J, BLUMENSCHEIN GR, ET AL. Intensive postoperative chemo-immunotherapy for patients with stage II and III breast cancer. Cancer 41:1064–1075, 1978.
2. GEHAN EA. Comparative clinical trials with historical controls: a statistician's view. Biomedicine 28:13–19, 1978.
3. GEHAN EA. Adjustment for prognostic factors in the analysis of clinical studies. *In* UICC Technical Reports Series 3. Methods and Impact of Controlled Therapeutic Trials in Cancer 36:35–74, 1978.
4. COX DR. Analysis of Binary Data. London, England, Methuen, 1970.
5. LEE ET. A computer program for linear logistic regression analysis. Computer Prog Biomed 4:80–92, 1974.
6. COX DR. Regression models and life tables. J R Stat Soc 34:187–220, 1972.
7. KAY R. Proportional hazard regression models and the analysis of censored survival data. Appl Sta. 26:227–237, 1977.
8. BRESLOW N. A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. Biometrika 57:579–594, 1970.
9. GEHAN EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203–224, 1965.
10. FISHER B, GLASS A, REDMOND C, ET AL. L-Phenylalanine mustard (L-PAM) in the management of primary breast cancer. An update of earlier findings and a comparison with dose utilizing L-PAM plus 5-fluorouracil (5-FU). Cancer 39:2883–2903, 1977.
11. BONADONNA G, ROSSI A, VALAGUSSA P, ET AL. The CMF program for operable breast cancer with positive axillary nodes. Updated analysis on disease-free interval, site of relapse and drug tolerance. Cancer 39:2904–2915, 1977.

# Use of Prognostic Factors in Improving the Design and Efficiency of Clinical Trials in Childhood Leukemia: Children's Cancer Study Group Report

**Denis R. Miller, Sanford Leikin, Vincent Albo, Leonard Vitale, Harland Sather, Peter Coccia, Mark Nesbit, Myron Karon,† and D. Hammond [1,*]**

## SUMMARY

**Prognostic factors are exerting an increasingly apparent and important influence on the results of comparative treatment regimens in childhood leukemia. Since February 1975, 883 previously untreated children with acute lymphocytic leukemia were entered in a protocol designed to evaluate prognostic factors that influence induction rate, remission duration, and survival and to identify subsets of patients with a high risk of early failure. Stratification was based upon the initial wbc count, and standard and more intensive induction and maintenance therapy programs were compared. The combination of older age (> 10 years), high initial wbc counts ($\geq 20 \times 10^9$/liter), and low serum IgG is associated with a significantly lower rate of induction of complete remission. The following factors were associated with an unfavorable prognosis with regard to bone marrow relapse and death: wbc count of $\geq 20 \times 10^9$/liter; age > 7 years old; CNS disease at diagnosis; $L_2$ lymphoblasts; hemoglobin level ≥ 11 g/dl; decreased concentration of serum immunoglobulin; and $M_3$ marrow on Day 14 of induction. None of the 41 patients in the low-risk prognostic group (3–7 years old with an initial wbc count of $< 10 \times 10^9$/liter) with $L_1$ blast cells, normal Ig, and $M_1$ marrow on Day 14 has sustained an adverse event. Identification of prognostic factors should permit more efficient and effective design of future protocols by refining the classification of acute lymphocytic leukemia, decreasing the toxic risk to patients with a better prognosis, and improving the relatively poor remission duration and survival for patients with a poor prognosis.**

[Cancer Treat Rep 64:381–392, 1980]

---

[1]This study was conducted by the Children's Cancer Study Group and was supported by Public Health Service grants (*CA*) from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare. The following institutions and investigators participated in the study: Children's Hospital National Medical Center, Washington, DC (*S. Leikin;* CA-03888); Children's Hospital of Los Angeles, CA (*G. Higgins* and *J. Ortega;* CA-02649); University of Southern California at Los Angeles (*D. Hammond, J. Weiner, R. Honour,* and *H. Sather;* CA-13539); Babies Hospital, New York, NY (*J. Wolff;* CA-03526); Children's Hospital of Pittsburgh, PA (*V. Albo;* CA-07439); Children's Hospital of Columbus, OH (*W. Newton* and *I. Ertel;* CA-03750); Children's Orthopedic Hospital, Seattle, WA (*R. Chard;* CA-10382); University of Minnesota, Minneapolis (*M. Nesbit;* CA-07306); Children's Memorial Hospital, Chicago, IL (*G. Honig;* CA-07431); University of Utah Medical Center, Salt Lake City (*E. Lahey;* CA-10198); University of Rochester, NY (*M. Klemperer;* CA-11174); Children's Hospital of Philadelphia, PA (*M. Donaldson;* CA-11796); Cornell Medical Center, New York, NY (*D. R. Miller;* CA-14557); Indiana University, Indianapolis (*R. Bachner;* CA-13809); New Jersey College of Medicine, Newark (*L. Vitale;* CA-12637); Harbor General Hospital, Torrance, CA (*J. Finklestein;* CA-14560); University of Michigan, Ann Arbor (*R. Heyn;* CA-02971); University of California, San Francisco (*A. Ablin;* CA-17829); Children's Hospital of Louisville, KY (*D. R. Kmetz*); University of British Columbia, Vancouver, Canada (*M. Teasdale;* Vancouver Foundation); Princess Margaret Hospital, Toronto, Ontario, Canada (*M. Sonley;* Ontario Cancer Treatment and Research Foundation grant 41-17); University of Wisconsin, Madison (*N. Shahidi;* CA-05436); University of Oregon, Portland (*R. Neerhout*); and Rainbow Babies and Children's Hospital and Case Western Reserve University, Cleveland, OH (*S. Gross*).

†*Deceased.*

**Reprint requests to:* Children's Cancer Study Group, Operations Office, 1721 N Griffin Ave, Los Angeles, CA 90031.

During the past 15 years, a more rational approach to the treatment of childhood acute lymphocytic leukemia (ALL) has been responsible for the observation that > 50% of children with ALL are alive ≥ 5 years after diagnosis. This successful strategy (1,2) has recognized the importance of (*a*) promptly achieving a complete remission, (*b*) using multiple-drug combinations in remission induction and maintenance therapy, (*c*) providing intensive supportive care during remission induction and maintenance therapy, and (*d*) protecting the CNS and possible other sanctuaries. An important objective of future studies should be to tailor therapy to the patient and the prognostic factors. In this way, the toxic risks to patients with favorable prognostic factors can be reduced.

The 5-year survival rate of patients treated by the Children's Cancer Study Group (CCSG) since 1969 has increased from < 20% in studies begun ≥ 10 years ago (3–5) to > 60% in more recent studies (6,7). CNS prophylaxis either as cranial irradiation with intrathecal methotrexate (MTX), craniospinal irradiation (1,2), or intrathecal MTX alone (8,9) has been a key element in the improved prognosis of childhood leukemia. The purpose of this present paper is to review previous studies performed by the CCSG which relate to prognostic factors and to report preliminary results of an ongoing study designed to evaluate prognostic factors.

## CCSG-803

The role of prognostic factors in determining the results of a treatment program has recently come into sharper focus (10). In the CCSG-803 study (4), the effect of the additives actinomycin D, mechlorethamine, and 5-fluorouracil on a basic maintenance therapy program consisting of 6-mercaptopurine (6-MP) alternating with MTX was evaluated. When remission duration was analyzed without regard to prognostic factors, no significant differences between any of six treatment regimens were noted. However, in patients with an initial wbc count of $< 20 \times 10^9$/liter, a statistically significant beneficial effect of additive therapy was observed in three regimens, whereas no beneficial effect of additive therapy was observed in patients with an initial wbc count of $\geq 20 \times 10^9$/liter.

Although CCSG-803 did not provide a significant breakthrough in induction or maintenance therapy, the interaction between strata based upon the initial wbc count and treatment regimens was an important finding and confirmed the experience of the Medical Research Council Concord trial in which an interaction effect between MTX in maintenance therapy and low initial wbc count was observed (11). Although several prognostic factors in ALL had been previously identified, these reports were the first to demonstrate the importance of prognostic factors in evaluating the results of a comparative clinical trial. With the overall improvement in induction rates to > 90% and the prolongation of remission beyond 3 or 4 years, the importance of front-end prognostic factors takes on added significance, now that a definitive answer to a clinical trial in childhood leukemia may require 7–10 years to complete. An accurate knowledge of host-disease interaction should provide front-end prognostic criteria that can be used to stratify the patient population and to increase the efficiency of the clinical trial (12).

## CCSG-101

The CCSG-101 study was designed to evaluate CNS prophylaxis. Without regard to prognostic factors, all patients were given prednisone (PDN), vincristine (VCR), and L-asparaginase (L-ASP), and were then randomized to receive one of four different programs of CNS prophylaxis: (*a*) craniospinal radiation plus extended-field irradiation to nodes, liver, spleen, kidneys, and gonads; (*b*) craniospinal radiation alone; (*c*) cranial radiation plus intrathecal MTX; or (*d*) intrathecal MTX alone. All patients received similar maintenance therapy consisting of daily 6-MP, weekly MTX, and monthly pulse doses of PDN and VCR.

Three distinct prognostic groups, based upon wbc count at diagnosis and age, were identified retrospec-

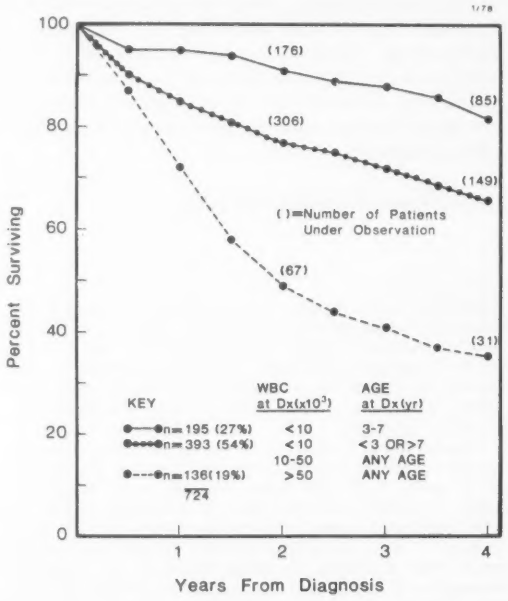



FIGURE 1.—Survival of children with ALL entered in CCSG-101 analyzed by prognostic groups.

tively. Survival in the three groups is shown in figure 1 (13). One hundred and ninety-five (27%) of the 724 evaluable patients entered in the study had an initial wbc count of $< 10 \times 10^9$/liter and were 3–7 years old. Eighty-five percent of the patients in this "low-risk" group are alive 4 years after diagnosis. A second group with a survival curve similar to that of all patients entered in the study made up 54% of the total study population. Children in this "average-risk" group either had an initial wbc count of $< 10 \times 10^9$/liter and were $< 3$ or $> 7$ years old, or had an initial wbc count of between 10 and $50 \times 10^9$/liter and were any age $< 15$ years old; 70% of the patients in this group are alive at 4 years. A third distinct group with a "high-risk" of early relapse consisted of 135 patients (19% of the total No. of patients entered) who had an initial wbc count of $> 50 \times 10^9$/liter. Fifty percent of the patients in this high-risk group were dead at 24 months, and < 40% are alive at 4 years. Intrathecal MTX alone was inferior to CNS prophylaxis employing radiation in patients with high ($\geq 20 \times 10^9$/liter) and low ($< 20 \times 10^9$/liter) initial wbc counts.

Using the Cox regression model (14), we determined the weighted contribution of a number of individual prognostic factors to the prediction of remission duration. The unfavorable variables, wbc count ($> 20 \times 10^9$/liter), age ($< 3$ or $> 7$ years), marked lymphadenopathy, hemoglobin level ($\geq 11$ g/dl), sex (male), and platelet count ($< 100 \times 10^9$/liter), were significant factors and contributed 96% to the prediction of the duration of remission (table 1).

## CCSG-141

CCSG-141 was a prospective study of prognostic factors in ALL. The prognostic factors described in CCSG-803 and CCSG-101 were determined by retrospective analysis of these studies. CCSG-141 was initiated in February 1975 primarily to establish the weight of front-end prognostic factors that could be used to stratify a patient population and to identify subsets of patients with a high risk of early relapse and death. A secondary objective was to develop more effective therapy for patients with high initial wbc counts ($\geq 20 \times 10^9$/liter) whose prognosis is poor.

The schema of CCSG-141 is presented in figure 2. Patients with an initial wbc count of $< 20 \times 10^9$/liter were given induction therapy of PDN, VCR, and L-ASP, cranial radiation therapy of 2400 rads plus six weekly injections of intrathecal MTX, and maintenance therapy consisting of daily 6-MP, weekly MTX, and monthly VCR and PDN until 12 months, at which time VCR and PDN were discontinued (low regimen 1). Children with an initial wbc count of $\geq 20 \times 10^9$/liter were randomized to receive either regimen 1 or 2, consisting of PDN, VCR, L-ASP, and cyclophosphamide induction therapy, more intensive CNS prophylaxis with a higher dose of 6-MP and intrathecal MTX given twice weekly, and intermittent cycles of POCA (PDN, VCR, cytosine arabinoside, and Adriamycin) or POMP (PDN, VCR, MTX, and 6-MP) until 12 months. Thereafter, maintenance therapy was similar to regimen 1.

The following factors were evaluated to determine the weighted contribution of each factor to prognosis: (*a*) leukemia thrust or virulence as measured by routine hematologic studies and physical measurement of mass disease and organomegaly; (*b*) cytomorphology as defined by the French-American-British collaborative study (15); (*c*) PAS histochemistry as quantified by the method of Hayhoe et al (16) (both the total PAS score and the percentage of PAS-positive blast cells were measured); (*d*) rate of remission induction by the performance of serial bone marrows on Days 0, 14, and 28 (17); and (*e*) host factors including sex, race, HLA typing (18), humoral immunity (quantitative immunoglobulins) (19), and cell-mediated immunity (as measured by skin test response to recall antigens [20] and serum inhibitors of blastogenesis in vitro) (21). Life-table curves were analyzed by the methods of Peto et al (22) and Mantel (23). The study was closed to patient entry in March 1977 after 883 patients had been entered. Although the results of remission duration and survival are still preliminary, subsets

TABLE 1.—CCSG-101 marrow remission duration

| Prognostic factor | Increment in log likelihood | Contribution to prediction (%) | Significance (*P* value) |
|---|---|---|---|
| Wbc count | 19.441 | 38.76 | < 0.000001 |
| Nodes | 9.512 | 18.96 | 0.000074 |
| Age (age)$^2$ | 8.885 | 17.71 | 0.00014 |
| Hemoglobin level | 4.173 | 8.32 | 0.0039 |
| Sex (male) | 3.536 | 7.05 | 0.0078 |
| Platelet count | 2.631 | 5.24 | 0.0220 |
| Spleen | 0.893 | 1.78 | 0.1814 |
| Race | 0.582 | 1.16 | 0.2806 |
| Liver | 0.503 | 1.00 | 0.316 |
| Mediastinal mass | 0.007 | 0.01 | 0.906 |

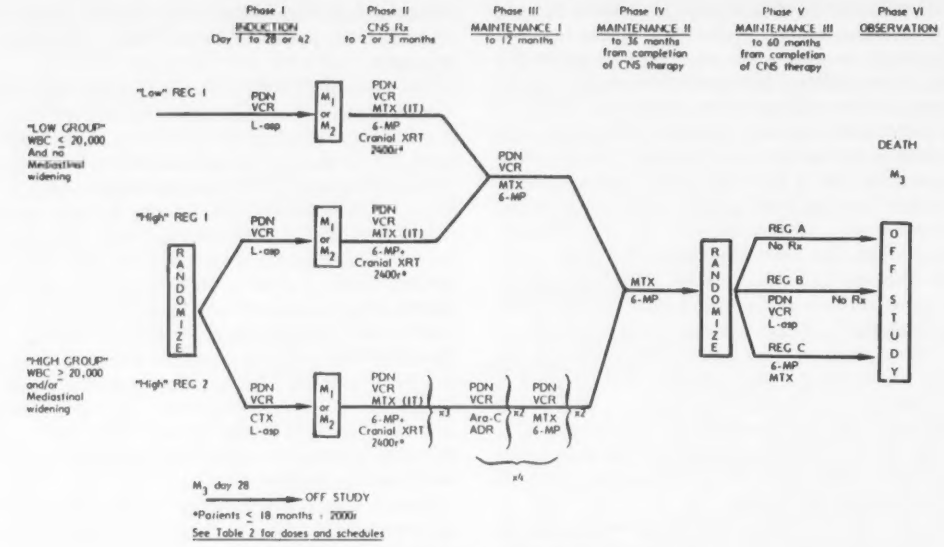



FIGURE 2.—Schema for CCSG-141.

of patients at a high risk of early relapse and of patients with a particularly good prognosis have been identified.

## Results of CCSG-141

### *Induction*

The overall complete remission (CR) rate was 94.2%; 95% in low regimen 1, 93% in high regimen 1, and 91% in regimen 2. Key factors predictive of remission induction were age, initial wbc count, and serum immunoglobulin (Ig). The combination of age > 10 years, wbc count ≥ 20 × $10^9$/liter, and decreased serum IgG (2 SDs below age-corrected mean) was associated with a significantly lower CR rate than any other subset (table 2). At one end of the spectrum, patients < 10 years old with initial wbc counts of < 20 × $10^9$/liter and normal or increased IgG had a CR rate of 98.3%; in contrast, children > 10 years old with initial wbc counts of ≥ 20 × $10^9$/liter and depressed IgG levels had a CR rate of only 55.6% ($P$ = 0.03). Multivariate analysis showed that IgG had a significant impact on the likelihood of successful remission induction, after adjusting for the association of IgG with other prognostic factors.

Patients with an $M_3$ (> 25% lymphoblasts) marrow rating on Day 14 had a significantly lower CR rate (78.1%) than children with either an $M_1$ (97.4%) or an $M_2$ (93.8%) marrow rating on Day 14 (table 3).

Other factors associated with a lower CR rate and a higher rate of induction failure included Down's syn-

TABLE 2.—CCSG-141: induction success ($M_1$ marrow) by age-wbc count-IgG categories

| Marrow status | | Age-wbc count category | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Normal Ig | | | | Decreased Ig | | | |
| | | LL* | LH* | HL† | HH† | LL* | LH* | HL† | HH† |
| $M_1$ | No. of patients (%) | 237 (98.3) | 101 (97.1) | 61 (82.4) | 29 (87.9) | 44 (100) | 24 (88.9) | 10 (100) | 5 (55.6) |
| Not $M_1$ | No. of patients | 4 | 3 | 5 | 4 | 0 | 3 | 0 | 4 |
| Total | No. of patients | 241 | 104 | 66 | 33 | 44 | 27 | 10 | 9 |
| | $P$ value‡ | NS | NS | NS | NS | NS | 0.07 | NS | 0.03 |

*LL = age < 10 yrs, wbc count < 20 × $10^9$/liter; LH = age < 10 yrs, wbc count > 20 × $10^9$/liter.
†HH = age > 10 yrs, wbc count > 20 × $10^9$/liter; HL = age > 10 yrs, wbc count < 20 × $10^9$/liter.
‡NS = not significant.

TABLE 3.—Induction failures in CCSG-141

| Factor | No. of induction failures/total patients (%) | Observed/expected failure rate |
|---|---|---|
| $L_3$ morphology | 2/6 (33.3) | 5.74 |
| $M_3$ marrow, Day 14 | 16/73 (21.9) | 3.77 |
| CNS disease, at diagnosis | 5/30 (16.7) | 2.88 |
| < 40% PAS-positive lymphoblasts | 21/154 (13.6) | 2.34 |
| Down's syndrome | 2/15 (13.3) | 2.29 |
| PAS score < 100 | 22/178 (12.4) | 2.14 |
| All patients | 51/883 (5.8) | 1.00 |

drome (13.3% of induction failures), CNS disease at diagnosis (16.7%), $L_3$ lymphoblast morphology (33.3%), PAS score of < 100 (12.4%), or < 40% PAS-positive lymphoblasts (13.6%). Sex, race, mediastinal mass, hemoglobin level, platelet count, and marrow cellularity were not predictive of induction therapy success or failure. Patients with $L_3$ lymphoblast morphology had a 5.7-fold higher failure rate. Although not significant ($P = > 0.5$), the induction therapy failure rate in the $L_2$ morphology group was 1.5 times higher than that in the $L_1$ group.

## EFFECT OF PROGNOSTIC FACTORS ON REMISSION DURATION AND SURVIVAL

### Lymphoblast Morphology

A total of 536 evaluable Wright and/or Wright-Giemsa-stained slides were reviewed by two independent observers of which 85% were classified as "$L_1$" ALL. $L_1$ blast cells have scanty cytoplasm, clumped or fine homogenous chromatin, and 0–2 indistinct nucleoli (fig 3a); 14% were classified as "$L_2$" ALL. $L_2$ blast cells are larger than $L_1$ blast cells, have moderate to abundant cytoplasm, fine homogenous, unclumped nuclear chromatin, and one to three distinct nucleoli (fig 3b). Slides with a mixed pattern ($L_1$–$L_2$) were classified as $L_2$ (fig 3c); 1% of the patients had "$L_3$" blast cells that are characteristic of Burkitt's leukemia with homogenous chromatin, intensely blue cytoplasm, and marked vacuolization.

We were unable to classify leukemic lymphoblasts using the criteria of Mathé et al (24) because of considerable overlap among their four categories. Nor did we (and others [25]) find that blast cell size (26) alone was helpful prognostically. Lee and co-workers (27) scored lymphoblasts according to cell size and nucleoli and found that smaller blast cell size and the absence of nucleoli were associated with a favorable prognosis. We found a correlation between $L_1$, $L_2$, and $L_3$ lymphoblast morphology and a total score based upon amount of cytoplasm, number of nucleoli, and vacuolization; the score of $L_1$ blast cells was lowest, $L_2$ was intermediate, and $L_3$ was highest.

$L_1$ lymphoblast morphology is associated with a significantly lower rate of bone marrow relapse than $L_2$ or $L_1$–$L_2$ blast cells when all patients are combined (table 4, fig 4), in patients with initial wbc counts of $< 10 \times 10^9$/liter (fig 5), and in the average-risk prognostic group (fig 6) (22). Morphology was not associated with a higher relapse rate in patients in the high-risk prognostic group.

### Day-14 Bone Marrow

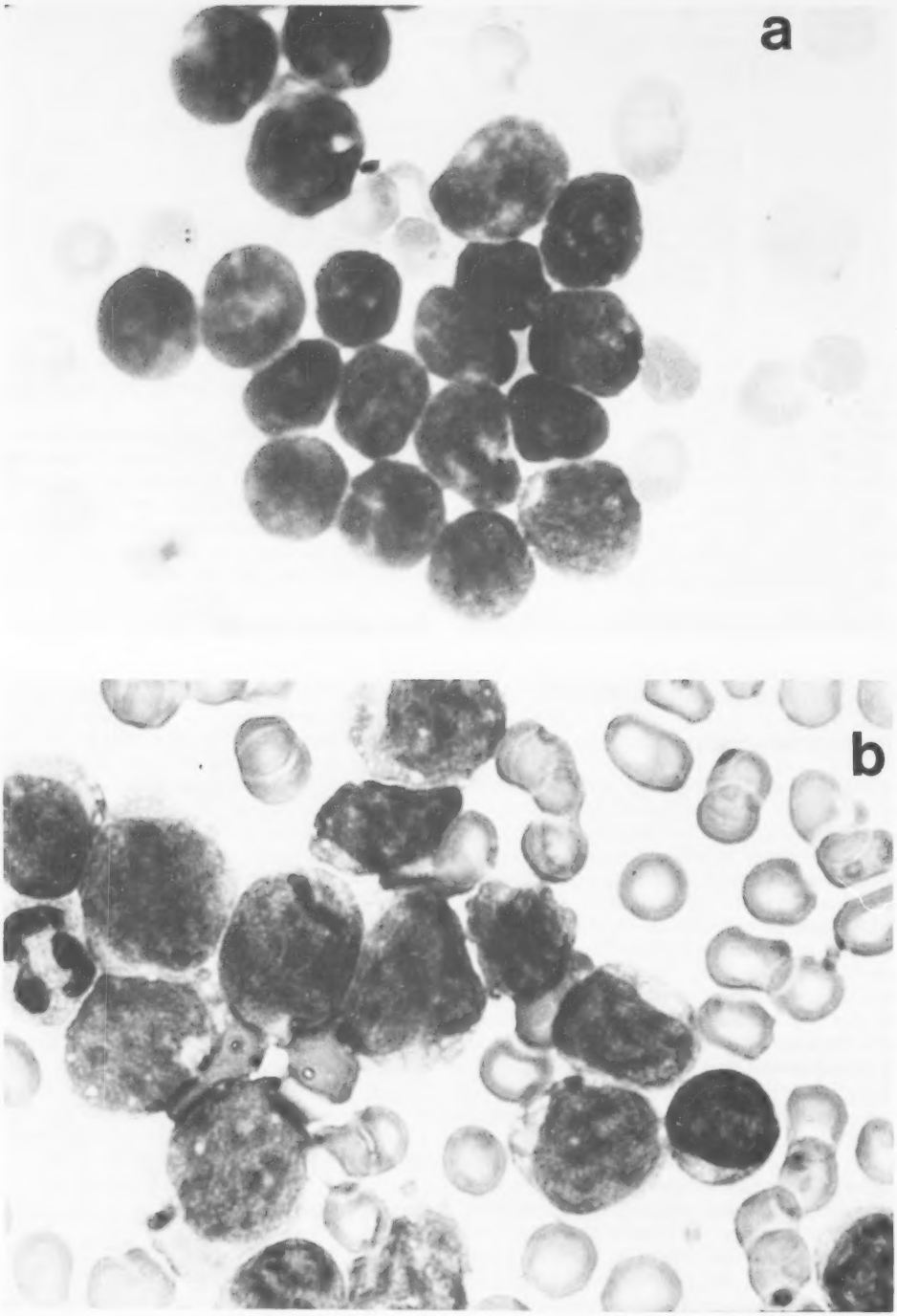
Patients with $M_1$ marrow on Day 14 have a lower relapse rate than patients with an $M_2$ or $M_3$ marrow who achieved $M_1$ by Day 28 or 42. This is true for all patients and for patients with initial wbc counts of $< 20 \times 10^9$/liter. For children with an initial wbc count of $> 20 \times 10^9$/liter, the effect of Day-14 marrow is less pronounced ($P = 0.07$) than in the low wbc count group, although still suggestive of an effect. The life-table adjusted relative relapse rate shows that the high wbc group has a 1.8-fold higher occurrence of relapse in the group with $M_2$ or $M_3$ marrow on Day 14. Thus, early response to induction therapy is a significant factor, confirming the findings of Jacquillat et al (17).

### Immunoglobulin Levels

As with induction, Ig levels are related to remission duration and relapse rates. The relapse rate is higher in patients with depressed levels of two of the three classes of Ig than in patients with normal levels. The differences are most striking in patients with decreased levels of all three Ig (fig 7).

### Other Prognostic Factors

Additional factors in CCSG-141 associated with a significantly higher relapse rate included initial wbc counts $\geq 20 \times 10^9$/liter, age > 7 years, CNS disease at diagnosis,


a
b

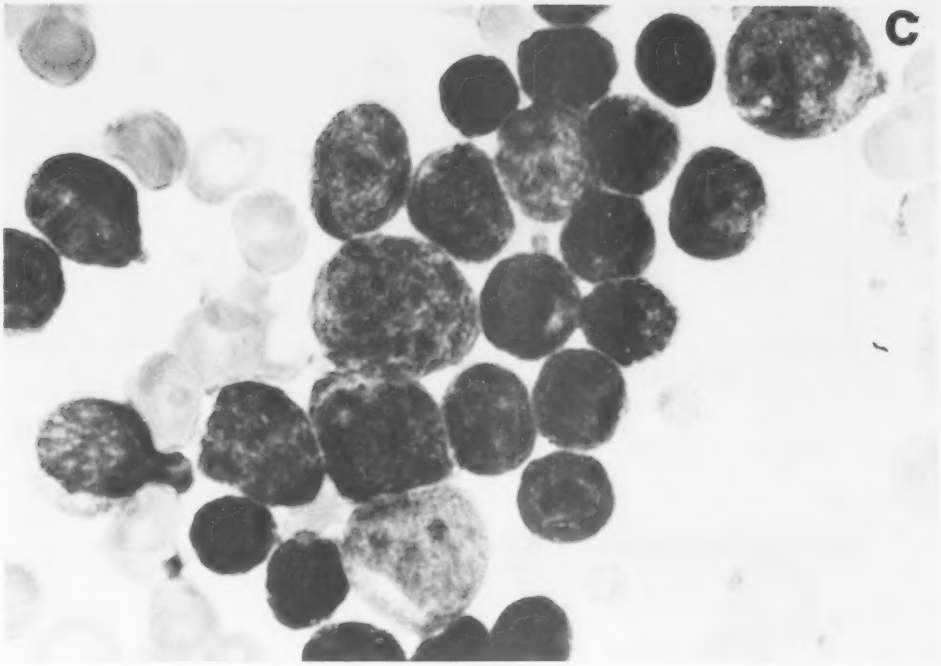



FIGURE 3.—Lymphoblast morphology: *a*. $L_1$ lymphoblasts; *b*. $L_2$ lymphoblasts; *c*. $L_1/L_2$ lymphoblasts.

TABLE 4.—Morphologic subgroups in CCSG-141

| Subgroup | No. of patients (%) | Observed/expected relapse rate |
|---|---|---|
| $L_1$ | 458 (84.3) | 0.80 |
| $L_1/L_2$ | 59 (10.9) | 1.92 |
| $L_2$ | 19 (3.5) | 4.14 |
| $L_3$ | 7 (1.3) | Not analyzed |
| Totals | 543 (100) | |

hemoglobin level ≥ 11 g/dl, and black race. Sex, HLA type, and skin test response to recall antigens were not predictive of relapse.

### Re-definition of Prognostic Groups

Using the previously defined prognostic groups based upon age and wbc count at diagnosis (13), life-table curves of time to marrow relapse in CCSG-141 are similar to those observed in CCSG-101. Complete data on Ig, Day-14 bone marrow, and blast cell morphology were available for 362 patients. Patients in prognostic groups were further subclassified as "good" if Ig levels were normal or elevated, if Day-14 marrow was $M_1$, and if blast cell morphology was $L_1$, and "poor" if one or more Ig levels were depressed, *or* if Day-14 marrow was $M_2$ or $M_3$, *or* if morphology was $L_2$ or $L_3$ (fig 8). All low-risk "good" group patients continue in CR and have not sustained any adverse events (relapse or death). Furthermore, patients with "good" factors within the average- and high-risk groups have lower relapse rates than patients with "poor" prognostic factors. Thus, Ig, Day-14 marrow, and blast cell morphology are additional prognostic factors that further refine prognostic groups previously delineated by age and wbc count alone.

## DISCUSSION

The rationale for front-end stratification was to improve the efficiency of the comparative trial by reducing the effect of host and disease variability, to tailor therapy by reducing the toxic risk to patients with a better prognosis, and to identify those patients who need more intensive therapy. Traditionally, therapy has been inten-

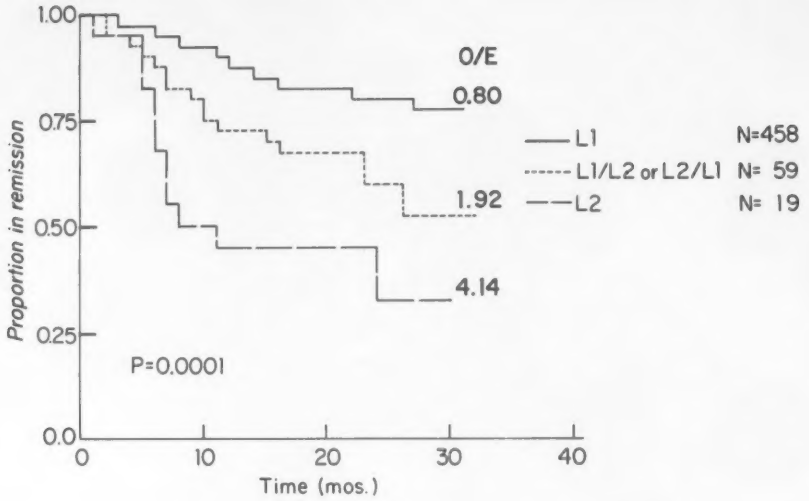



FIGURE 4.—Remission duration from start of CNS therapy in CCSG-141 by cell morphology. O/E = ratio of observed/expected relapses.

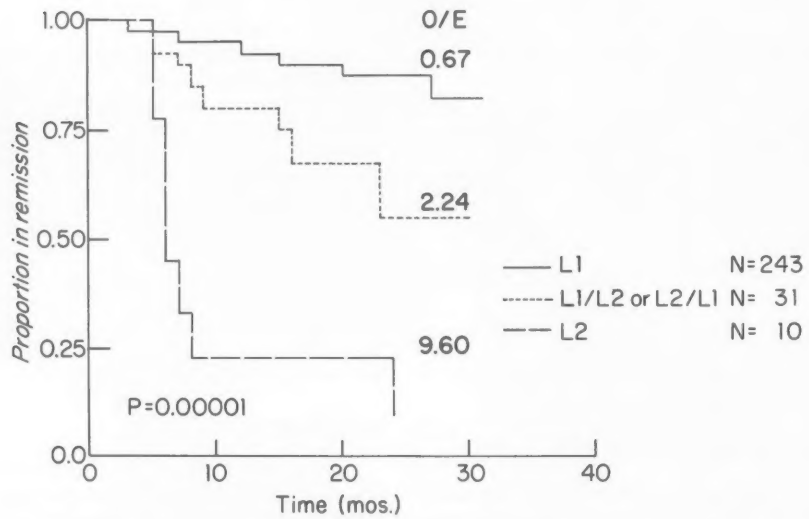



FIGURE 5.—Remission duration from start of CNS therapy in CCSG-141 by cell morphology in patients with initial wbc counts of $< 10 \times 10^9$/liter.

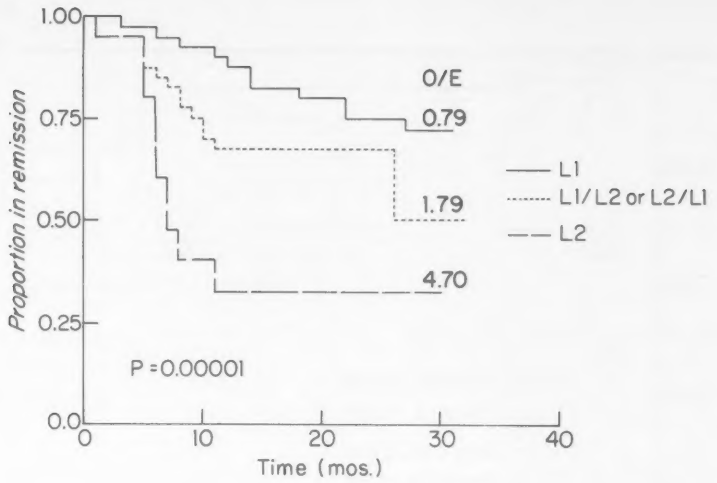



FIGURE 6.—Remission duration from start of CNS therapy in CCSG-141 by cell morphology in average-risk prognostic group.

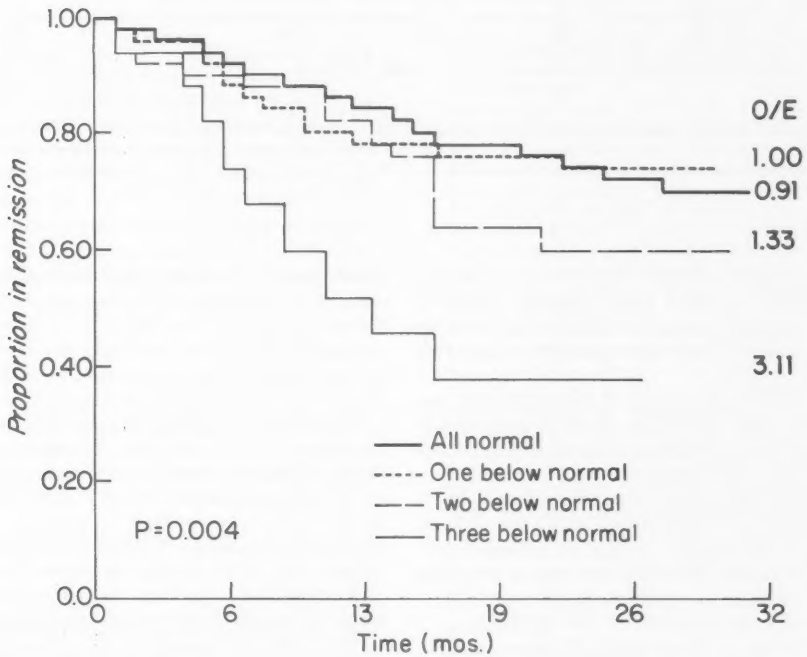



FIGURE 7.—Relapse rate in patients entered in CCSG-141 with normal or ≥ 1 depressed levels of Ig. Respective N values are: all normal, 488; 1 below normal, 130; 2 below normal, 54; all depressed, 17.

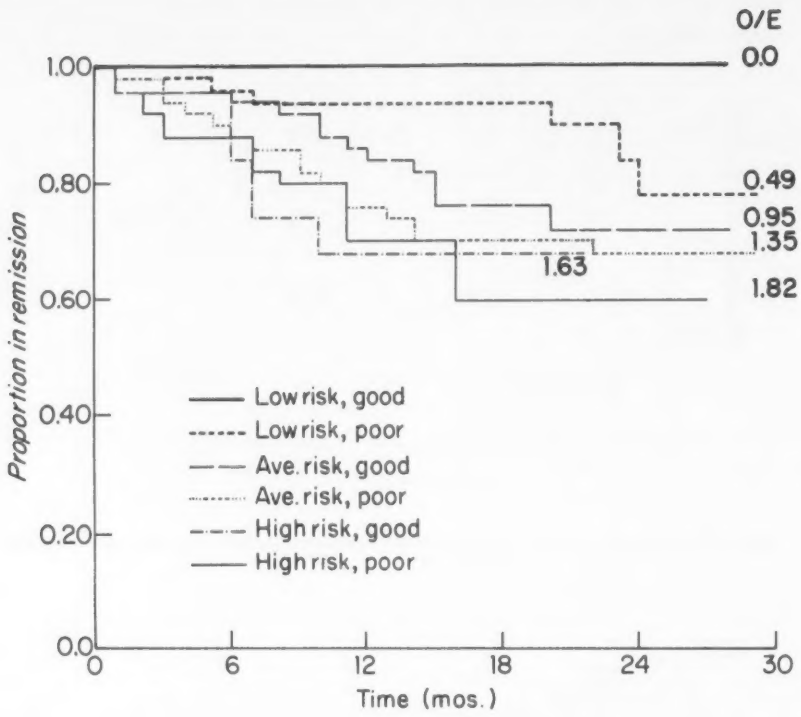



FIGURE 8.—Time to 1st event (marrow relapse, CNS relapse, or death) in patients with complete data for Ig, Day-14 marrow status, and blast cell morphology by prognostic groups. "Good" subset had normal or elevated Ig, $M_1$ marrow on Day 14, and $L_1$ morphology. "Poor" subset had depressed Ig, $M_2$ *or* $M_3$ marrow on Day 14, *or* $L_1/L_2$, $L_2$, *or* $L_3$ morphology. No events have occurred in the low-risk good group.

sified in all patients with ALL regardless of prognostic factors. Although remission rates, remission duration, and survival have improved, an escalation of the intensity and toxicity of therapy may not be justified in good-risk patients.

An analysis of a single institution, nonrandomized, single-limb study points out a dilemma of the uncontrolled trial that relies on historical controls. The L-2 protocol of the Memorial Sloan-Kettering Cancer Center (New York, NY) (9) employed intensive chemotherapy in 75 previously untreated children with ALL. The 60% remission rate at 48 months is virtually identical to that obtained in CCSG-101. When L-2 is evaluated according to prognostic groups, an ordering effect is seen that is similar to the results obtained in CCSG-101 (fig 9).

In the L-2 protocol, the 48-month CR rates are 90%, 65%, and 40% for low-, average-, and high-risk groups respectively. Comparable rates in the CCSG-101 study are 76%, 66%, and 45%, a clear demonstration that prognostic groups, not therapy, may determine remission success. The philosophy that more aggressive therapy is better therapy may be fallacious if the contributions of prognostic factors are not considered. Currently employed intensive regimens have not resulted in a significant improvement in the survival of poor-risk patients.

In summary, a number of prognostic factors must be considered in designing and evaluating treatment protocols. Favorable factors include initial wbc count $< 10 \times 10^9$/liter, age between 3 and 7 years, no CNS disease at diagnosis, $L_1$ lymphoblast morphology, normal Ig levels, hemoglobin level < 7 g/dl, and $M_1$ marrow on Day 14 of induction. Other studies have shown that white race, null cell markers (28), and present or increased glucocorticoid receptors (29) are favorable prognostic factors.

The CCSG has recently activated three different protocols for previously untreated patients with ALL. An initial stratification is based upon low-, average-, or high-risk prognostic grouping. Questions of either decreasing

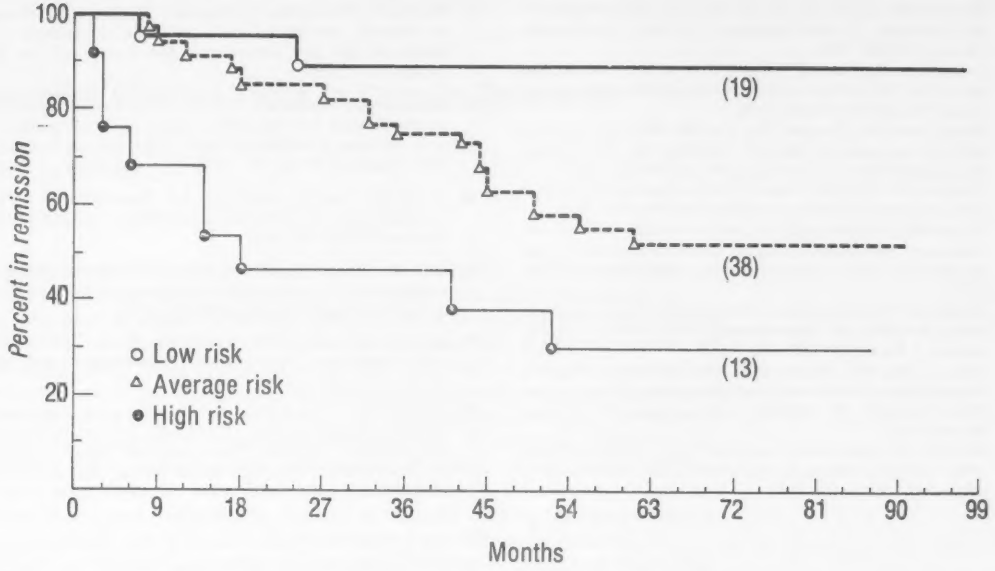



FIGURE 9.—Remission duration in patients treated on L-2 protocol at the Memorial Sloan-Kettering Cancer Center analyzed according to prognostic groups (personal communication and ref 9).

or increasing therapy will be asked in relatively homogenous groups of patients refined so that prognostic factors themselves will not influence the evaluation of the experimental treatment regimens. The cooperative group approach will also permit the entry of a large number of patients with statistically valid results anticipated within a reasonable period of time. A reclassification of childhood leukemia based upon cytomorphologic, histochemical, biochemical, and immunologic markers should further improve the encouraging treatment results now emerging.

## REFERENCES


1. MAUER AM, and SIMONE JV. The current status of treatment of childhood acute lymphoblastic leukemia. Cancer Treat Rev 3:17–41, 1976.
2. MAUER AM, SIMONE JV, and PRATT CB. Current progress in the treatment of the child with cancer. J Pediatr 91:523–539, 1977.
3. LEIKIN S, BRUBAKER C, HARTMANN J, ET AL. The use of combination therapy in leukemia remission. Cancer 24:427–432, 1969.
4. MILLER DR, SONLEY M, KARON M, ET AL. Additive therapy in the maintenance of remission in acute lymphoblastic leukemia of childhood: the effect of the initial leukocyte count. Cancer 34:508–517, 1974.
5. HEYN R, BORGES W, JOO P, ET AL. BCG in the treatment of acute lymphocytic leukemia. Blood 46:431–442, 1975.
6. ORTEGA JA, NESBIT ME, JR, DONALDSON MH, ET AL. L-Asparaginase, vincristine, and prednisone for induction of first remission in acute lymphocytic leukemia. Cancer Res 37:535–540, 1977.
7. MILLER DR, LEIKIN S, ALVO V, ET AL. New prognostic factors in childhood leukemia (ALL). Pediatr Res 12:469, 1978.
8. WANG JJ, FREEMAN AI, and SINKS LF. Treatment of acute lymphocytic leukemia by high-dose intravenous methotrexate. Cancer Res 36:1441–1444, 1976.
9. HAGHBIN M, TAN CC, CLARKSON BD, ET AL. Intensive chemotherapy in children with acute lymphoblastic leukemia (L-2 protocol). Cancer 33:1491–1498, 1974.
10. MILLER DR. Prognostic factors in childhood leukemia. J Pediatr 87:672–676, 1976.
11. Preliminary report to the Medical Research Council by the Leukemia Committee and the Working Party on Leukemia—Comparison of immunotherapy (B.C.G.), intermittent methotrexate, and no therapy after a five-month intensive cytotoxic regimen (Concord trial). Br Med J 4:189–194, 1971.
12. KARON M. Problems in evaluation of long-term results. *In* Recent Results in Cancer Research Nomenclature, Methodology and Results of Clinical Trials in Acute Leukemia (Mathé G, Pouillart P, and Schwarzenberg L, eds). New York, NY, Springer-Verlag, 1973, vol 43, pp 113–120.
13. COCCIA P, SATHER H, NESBIT M, ET AL. Interrelationship of initial wbc, age and sex in predicting prognosis in childhood acute lymphoblastic leukemia (ALL). Am Soc Hematol, 1976.
14. COX D. Regression models and life tables. J R Stat Soc B 34:187, 1972.
15. BENNETT JM, CATOVSKY D, DANIEL MT, ET AL. Proposals for the classification of the acute leukaemias. French-American-British (FAB) Co-operative Group. Br J Haematol 33:451–458, 1976.

16. Hayhoe FG, Quaglino D, and Doll R. The cytology and cytochemistry of acute leukaemia. A study of 140 cases. London, HMSO, 1964.
17. Jacquillat C, Weil M, Gemon MF, et al. Combination therapy in 130 patients with acute lymphoblastic leukemia. Cancer Res 33:3273–3284, 1973.
18. Rogentine GN, Trapani RJ, Yankee RA, et al. HL-A antigens and acute lymphocytic leukemia: the nature of the HL-A2 association. Tissue Antigens 3:470–476, 1973.
19. Khalifa AS, Take H, Cejka J, et al. Immunoglobulins in acute leukemia in children. J Pediatr 85:788–791, 1974.
20. Hersh EM, Whitecar JP, Jr, McCredie KB, et al. Chemotherapy, immunocompetence, immunosuppression and prognosis in acute leukemia. N Engl J Med 285:1211–1216, 1971.
21. Humphrey GB, Lankford J, Oleinick S, et al. Leukemic serum inhibition of phytohemagglutinin-induced transformation. J Immunol 113:63–69, 1974.
22. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer 35:1–39, 1977.
23. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 50:163–170, 1966.
24. Mathé G, Pouillart P, Sterescu M, et al. Subdivision of classical varieties of acute leukemia. Correlation with prognosis and cure expectancy. Rev Eur Etud Clin Biol 16:554–560, 1971.
25. Murphy SB, Borella L, Sen L, et al. Lack of correlation of lymphoblast cell size with presence of T-cell markers or with outcome in childhood acute lymphoblastic leukaemia. Br J Haematol 31:95–101, 1975.
26. Pantazopoulos N, and Sinks LF. Morphological criteria for prognostication of acute lymphoblastic leukaemia. Br J Haematol 27:25–30, 1974.
27. Lee SL, Kopel S, and Glidewell O. Cytomorphological determinants of prognosis in acute lymphoblastic leukemia of children. Semin Oncol 3:209–217, 1976.
28. Kersey JH, LeBien TW, Hurwitz R, et al. Childhood leukemia-lymphoma. Heterogeneity of phenotypes and prognoses. Am J Clin Pathol 72:746–752, 1980.
29. Konior GS, Lippman ME, Johnson GE, et al. Correlation of glucocorticoid receptor (GR) levels and complete remission duration (CRD) in "poor prognosis" acute lymphatic leukemia (ALL). Proc Am Assoc Cancer Res and ASCO 18:353, 1977.

# Sequential Clinical Trials in Cancer Research

Stephen L. George [1,*]

## SUMMARY

**Decisions to terminate clinical trials have been the subject of intense debate, both before and after the actual decisions are made. This situation reflects the unique character of studies involving human subjects in that the decision to stop is largely an ethical one in which the statistical considerations play a major, but not a dominating, role. This can be contrasted with, say, that of acceptance sampling in which statistical qua economic considerations completely determine the stopping rule. It is important for all investigators involved in clinical studies to be aware of the sequential nature of the analyses of those studies. In particular, they must be conscious of the inappropriateness of the usual "*P* values" when repeated analyses of accumulating data are done. There is a pressing need to develop more realistic statistical theory and methodology, as well as to apply existing techniques, where appropriate, particularly in view of the numerous practical and theoretic difficulties that hinder application of sequential techniques in clinical trials. This paper gives some historical background on sequential trials, describes the different classes of trials, and lists some important reasons why sequential trials are not used more often.**

[Cancer Treat Rep 64:393–397, 1980]

In a broad sense, all experimentation is sequential since experiments are planned and carried out in light of knowledge gained through past experiments and other information available to the experimenter. Thus, sequences of different experiments can be regarded as a sequential process whereby after each experiment the research worker chooses the next type and size of experiment based on current and previous results. This idea of a chain of experiments can be considered as a forerunner of sequential analysis (1) and is an essential part of the theory that has come to be known as the sequential design of experiments (2).

However, "sequential analysis," as it has been ordinarily interpreted in the statistical literature, refers to a general class of procedures for analyzing a body of data as it is accumulating with the intent to terminate the data-collection process as soon as some conclusion can be reached. The amount of data to be collected is not fixed in advance but is allowed to depend on the accumulating information. The designation "sequential" is used to distinguish this type of analysis from the classic (fixed-sample) statistical analysis in which no formal decision, action, or inference is taken until some predetermined amount of data has been collected, no matter what the early trends are. A key advantage cited by its proponents is that sequential procedures require, on the average, smaller numbers of observations than comparable non-sequential procedures.

In its most specific and restricted form, a sequential experiment is one in which observations are made one at a time, and after each observation a formal decision is taken either to take another observation or to terminate the procedure. The decision to stop is made only if the data are adequate to support some conclusion (eg, with respect to a specific hypothesis or with respect to the attained precision in an estimate) or permit some action to be taken. In a slightly more general form, a sequential experiment may be a multi-stage procedure in which each stage consists of taking one or more observations. The decision to continue to the next stage and the number of observations to be taken in the next stage are based on the cumulative results through the current stage.

[1]Biostatistics Section, St. Jude Children's Research Hospital, Memphis, TN.

**Reprint requests to:* Stephen L. George, PhD, Biostatistics Section, St. Jude Children's Research Hospital, 332 N Lauderdale, PO Box 318, Memphis, TN 38101.

In this paper the designation "sequential" will be applied to any procedure in which the accumulating data are "analyzed" in some way while the trial is in progress. This deliberately vague definition is used since the problems involved in sequential procedures are similar regardless of the use to which the analysis is put. Topics such as interim analyses of clinical trials and data-dependent or "adaptive" treatment allocation rules (3–5) are included in the class of sequential procedures even if no explicit decision to continue or discontinue is intended.

## HISTORY

The earliest formal work leading to the modern theory of sequential analysis was the two-stage or doubling sampling schemes of Dodge and Romig (6), in which the size of a second sample, including the possibility of taking no second sample, depends on the results in a preliminary sample. Bertky (7) extended the principle to multiple sampling for binomial distributions. All of these schemes resulted in expected sample sizes smaller than single-stage procedures with similar error probabilities. These early sequential schemes were set in the context of industrial acceptance sampling (ie, probability sampling of batches of manufactured items as part of quality control procedures) rather than in the context of scientific experimentation.

The major impetus to further development of sequential analysis came during and immediately after World War II through the work of G. A. Barnard in England and, especially, the work of A. Wald in the US. As in the earlier double- and multiple-stage procedures, the major applications of sequential analysis continued to be in the area of acceptance sampling. Wald's work, which was begun in 1943, was classified as restricted material under the Espionage Act during the war. After removal of the restricted classification, *Sequential Analysis*, the seminal work in this field, was published in 1947 (8). This book describes the sequential probability ratio test, and makes the conjecture, later proved (9), that of all tests with given control of error, the sequential probability ratio test has the minimum expected sample size. Average savings of 40% in the number of observations are not unusual.

In the 1940s, concurrent with the rise of the theory of sequential analysis in statistics, the concepts and theory of statistical experimental design began to be applied to the problems of medicine through large-scale clinical trials (10). These two spheres of activity were quite separate in their early development, and the timing of their development appears to be somewhat coincidental. In research involving human subjects, the ethical requirement to give individual subjects the best known therapy, including therapy given as part of a clinical trial, compelled consideration of experimental designs in which economy of observation (ie, in clinical trials, no more patients than "necessary") was a primary consideration. Further impetus was provided by a corollary and often misunderstood need to avoid premature stopping (ie, stopping before any conclusion can be reached) and the reality of serial entry of patients over time.

An important original text on the application of sequential concepts to clinical trials, *Sequential Medical Trials*, appeared in 1960 (11). This text attracted considerable attention, in part because of a critical book review which challenged the validity (when applied to experimental data) of the concepts on which the results were based (12). The controversy, which still persists, concerns the very foundations of statistical inference rather than whether or not sequential clinical trials should be used. Sequential procedures merely sharpen the differences among alternative approaches.

One aspect of the dispute that is brought into sharp relief by sequential procedures is the "paradox" of optional (not "optimal") stopping (13,14). Briefly stated, the controversy centers on the extent to which the nature of the rule leading to the decision to stop gathering data should be allowed to influence the final inferences from those data. The issues raised by optional stopping are similar to those raised by problems of "multiplicity" or repeated analyses of the same data (15). To proponents of the likelihood principle, in which all results yielding the same likelihood should yield the same inferences, the stopping rule is irrelevant (16). It is not the intention of this paper to explore these controversies, except in a cursory fashion, but it should be noted that their existence has had an impact on the conduct and interpretation of results of clinical trials, both sequential and nonsequential.

The remainder of this paper will consider two principal types of sequential clinical trials (SCTs): ordinary SCTs and followup SCTs.

## ORDINARY SEQUENTIAL CLINICAL TRIALS

In this paper, SCTs for which the time to "response" on each subject is immediate or very short relative to the total time available for study are called ordinary SCTs. These trials derive most directly from the original developments in sequential analysis, for which the most theory has been developed and for which a substantial savings in the total number of patients treated may be obtained. Ordinary SCTs are applicable to screening studies, including preclinical animal studies and phase I–II drug trials, as well as to comparative trials involving binary response variables. The most complete reference for the theory and application of ordinary SCTs is by Armitage (11), in which results are obtained primarily for the comparison of two treatment groups. Successive

entrants to the trial are paired in order to eliminate the effect of any time trends in the response to treatment. The procedures labeled "repeated significance tests" (RSTs) are particularly instructive in clarifying the relationship between sequential tests of hypotheses and the repeated application of ordinary, fixed-sample, tests on accumulating data (ie, the relationship between sequential analysis and optional stopping). Briefly, an RST is a sequential test in which repeated tests are made as the data accumulate, each test conducted as if the trial were designed as a fixed-sample trial. The RST procedures control the overall significance level by reducing the nominal significance level used at each test.

## FOLLOWUP SEQUENTIAL CLINICAL TRIALS

In this paper, SCTs for which the time to response is not immediate are called "followup" SCTs. In general, the time to response varies from patient to patient, although it may be fixed in certain situations. These studies are obviously not well-suited for the classic sequential analysis since, if the average time to response is long relative to the time needed to accrue large numbers of patients, the sequential approach can be of little value in reducing the number of patients entered (11). The potential saving is largely in terms of followup time, although in certain cases the use of sequential procedures offers the potential for changing the therapy of the remaining patients. Obviously, analogs of ordinary sequential procedures may not exist for all types of followup trials, and, even if they do exist, may no longer be optimal.

Although it is possible, in theory, to consider a continuous monitoring of results over time, the practical problems of carrying out such an analysis are forbidding. Accordingly, sequential procedures for followup SCTs can be classified into two types.

1. *Event-sequential.*—Here, analyses are carried out after certain events have occurred, eg, after each failure or groups of failures. The time between analyses is a random variable. Most of the proposals for followup SCTs have been of the event-sequential type. In those cases in which the time to response is noninformative (ie, in which the delay is merely a "nuisance" variable independent of the eventual response), it is natural to employ such a procedure since information concerning the effects of treatment is available only at the times of failure, not between failures. When the followup time itself is the variable of interest, partial information on response is available on each subject in the form of time-censored observations. In this setting, Breslow (17) developed a test which compares two exponential distributions by deriving the analogy between a simple test statistic and ordinary Brownian motion (18), and Armitage (11) suggests the application of a nonparametric test, the log-rank test.

2. *Time-sequential.*—Analyses are carried out at certain times. The number of events between these times is a random variable. This type of procedure is most appropriate in certain cooperative clinical trials. Time-sequential procedures have rarely been applied to clinical trials because of technical difficulties. Recently, attempts to approach this problem in a formal way have begun to be reported in the literature. Canner (19) proposed a procedure for monitoring treatment differences between two treatments using a normalized test statistic and assuming underlying exponential distributions for the time to response. The power of the test keeping the overall type I error rate constant is developed by simulation studies. Practical application of such procedures is limited by explicit distributional assumptions and pre-specified (fixed) number of "looks" at the data, by the implicit assumption that there is no "lag" in reporting the data (ie, all data are in and current), and by the lack of adjustment for prognostic factors.

Many sequential procedures for followup SCTs have been covered under the heading of "early stopping" procedures (20–25). The trial is monitored in such a way that the length of followup is no longer than necessary, so that, as mentioned earlier, the potential savings in these trials are in length of study rather than in number of patients entered. Early stopping may be important because it allows for early reporting of results and for the investigator to change therapy in any patient receiving an "inferior" treatment. However, it must be noted that there is no guarantee that patients initially receiving an inferior treatment will benefit from the switch to the "superior" treatment, since the benefit may have been derived from early treatment. Indeed, switching may have a detrimental effect. In any case, the decision must be made with care since there will be no way to retrieve the information concerning longer-term results which is "lost" by early stopping.

## BARRIERS TO THE USE OF SEQUENTIAL CLINICAL TRIALS

Despite the appeal of sequential procedures in clinical trials and the extensive theory already developed, the use of explicitly sequential procedures in clinical trials is apparently quite low. Since the major compilations of active clinical trials do not state whether or not the trial is designed as an SCT, it is impossible to estimate the percentage of SCTs in practice. Nevertheless, it is safe to guess, on the basis of published reports of trials in cancer, that the percentage is quite low. On the other hand, it is also safe to guess that nearly all clinical trials in cancer have repeated, informal, "interim" analyses since the data are accruing before any "final" analysis. Indeed, as mentioned previously, the sponsoring agency usually requires these analyses as progress reports.

Why are SCTs not more widely used? There are many

reasons, both practical and theoretic. Several of these reasons (in no particular order of importance) follow.

A. Simple unfamiliarity with the techniques. It is possible for a trained statistician, for example, to have had little or no exposure to the theoretic or practical aspects of sequential trials. Clinicians are even less likely to be aware of the design philosophy of sequential trials.

B. Administrative difficulty in conducting sequential trials can be formidable compared to similar fixed-sample trials. This difficulty is particularly acute in cooperative clinical trials. Simon (3) discusses this point and others in the context of adaptive treatment assignment.

C. Sequential procedures are most useful when the difference among therapies is large, so that the trial can be stopped quickly. For most clinical trials in practice, the difference among therapies tends to be moderate, so there is no need to analyze the data too frequently (26).

D. Since the decision to stop a clinical trial must consider ethical issues that involve measures of benefit and risk that are inherently incommensurate but may be more important than scientific/statistical issues, many investigators fear that, because of their rigid framework, most sequential stopping rules may be applied too mechanistically.

E. The simplistic nature of many sequential procedures is disturbing to many investigators. Clinical trials are complex undertakings in which the purpose is seldom, if ever, a simple test of hypothesis. It is sometimes claimed that sequential analysis is appropriate to situations such as acceptance sampling, but its dependence on an overly simplified hypothesis, ie, decision between two alternatives, renders it inappropriate to scientific interpretation.

F. The theory is not developed well enough to allow for "generally accepted" procedures, particularly in followup studies. Even the well-developed models often seem entirely inappropriate for clinical research.

G. Repeated analyses of accumulating data can, if properly accounted for, actually lead to a longer trial, on the average, than if fewer analyses had been performed (15). Stated another way, sequential procedures are less likely than fixed-sample procedures to detect "significant" differences among treatments with the same number of patients.

H. Lags in reporting data, particularly in cooperative trials, mean that each decision is based on incomplete data (27). In this situation, experienced investigators tend to view all interim analyses with suspicion.

I. Most clinical trials have more than one outcome of interest (eg, toxicity, efficacy: short-range and long-range) that are not properly considered in sequential analysis (28).

J. Most standard sequential procedures are two-sample procedures, but many clinical trials involve more than two treatments.

K. The disputes over fundamental issues of statistical inference, often focusing on sequential procedures, have led many investigators to believe (erroneously) that sequential procedures are more controversial than more "standard" nonsequential procedures.

## REFERENCES


1. Hotelling H. Experimental determination of the maximum of a function. Ann Math Stat 12:20–45, 1941.
2. Chernoff H. Approaches in sequential design of experiments. *In* A Survey of Statistical Design and Linear Models (Srivastava JN, ed). Amsterdam, The Netherlands, Elsevier North-Holland Publishing Co, 1975, pp 67–90.
3. Simon RH. Adaptive treatment assignment methods and clinical trials. Biometrics 33:743–749, 1977.
4. Zelen M. Play the winner rule and the controlled clinical trial. JASA 64:131–146, 1969.
5. Weinstein MC. Allocation of subjects in medical experiments. N Engl J Med 291:1278–1285, 1974.
6. Dodge HF, and Romig HG. A method of sampling inspection. Bell System Tech J 8:613–631, 1929.
7. Bertky W. Multiple sampling with constant probability. Ann Math Stat 14:363–377, 1943.
8. Wald A. Sequential Analysis. New York, NY, John Wiley, 1947.
9. Wald A, and Wolfowitz J. Optimum character of the sequential probability ratio test. Ann Math Stat 19:326–339, 1948.
10. Hill AB. Statistical Methods in Clinical and Preventive Medicine. New York, NY, Oxford University Press, 1962.
11. Armitage P. Sequential Medical Trials, 2nd ed. New York, NY, John Wiley, 1975.
12. Anscombe FJ. Sequential medical trials. JASA 58:365–383, 1963.
13. Hacking I. Logic of Statistical Inference. Cambridge, MA, University Press, 1965.
14. Darling DA. The birth, growth and blossoming of sequential analysis. *In* On the History of Statistics and Probability (Owen DB, ed). New York, NY, Marcel Dekker, 1976.
15. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. Science 198:679–684, 1977.
16. Cornfield J. Recent methodological contributions to clinical trials. Am J Epidemiol 104:408–421, 1976.
17. Breslow N. On large sample sequential analysis with applications to survivorship data. J Appl Probability 61:261–274, 1969.
18. Breslow N, and Huag C. Sequential comparison of exponential survival curves. JASA 67:691–697, 1972.
19. Canner PL. Monitoring treatment differences in long-term clinical trials. Biometrics 33:603–615, 1977.
20. Muenz LR, Green SB, and Byar DP. Applications of the Mantel-Haenszel statistic to the comparison of survival distributions. Biometrics 33:617–626, 1977.
21. Alling D. Early decision in the Wilcoxon two-sample test. JASA 58:713–720, 1963.
22. Chun D. On an extreme rank sum test with early decision. JASA 60:859–863, 1965.
23. Halperin M, and Ware J. Early decision in a censored Wilcoxon two-sample test for accumulating survival data. JASA 69:414–422, 1974.
24. Chatterjee SK, and Sen PK. Nonparametric testing under progressive censoring. Calcutta Stat Assoc Bull 22:13–50, 1973.
25. Louis TA. Sequential allocation in clinical trials comparing two exponential survival curves. Biometrics 33:627–634, 1977.

26. PETO R, PIKE MC, ARMITAGE P, ET AL. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 34:585–612, 1976.
27. GEORGE SL. Practical problems in the design, conduct and analysis of cooperative clinical trials. *In* Proceedings of the IX International Biometric Conference, 1:227–244, 1976.
28. BYAR DP, SIMON RM, FRIEDEWALD WT, ET AL. Randomized clinical trials. Perspectives on some recent ideas. N Engl J Med 295:74–80, 1976.

# Application of Restricted Sequential Design in a Clinical Protocol[1]

**Elizabeth I. Thompson**[2,3,*]

## SUMMARY

**Restricted sequential design is an alternative to fixed-sample analysis as a statistical tool in clinical trials. This paper presents a specific example of the rationale which led to the choice of sequential design in a clinical protocol evaluating granulocyte transfusions in children with leukemia. The main advantage of sequential design is that fewer patients may be necessary to declare statistical significance. Its application is limited to trials where (*a*) the result of therapy is easily defined, (*b*) the result of therapy is discernible in a short time interval, and (*c*) one randomization is being tested, although patients may be stratified. The main disadvantage of sequential design is that if the therapies being tested are similar in efficacy, the trial may require more patients than fixed-sample analysis. This potential disadvantage may be acceptable when the concern in designing a trial is to evaluate as quickly as possible a reputedly superior therapy.**

[Cancer Treat Rep 64:399–403, 1980]

Restricted sequential design is a well-defined statistical tool which at present is rarely used in clinical trials. This method has several potential advantages compared to fixed-sample analysis, but relatively few clinicians are familiar with its practical applications and limitations. In 1975, we incorporated restricted sequential design in a protocol evaluating granulocyte transfusions in children with leukemia. The purpose of this paper is to describe the rationale which led to the choice of sequential design, and to provide a specific example of the application of sequential design in a clinical setting. For simplicity, the historical background given will be limited to that which was available in 1975 when the study was first developed. Since then, patient selection has been modified, but this modification has not changed the rationale for the statistical approach. It will be discussed after the background and summary of the protocol are presented.

In 1975, the appropriate use of granulocyte transfusions in the supportive care of neutropenic patients with bacterial infections was even less well-defined than it is today. Reported patient series were often (*a*) uncontrolled, (*b*) included mainly adult patients with differing primary disorders, (*c*) used widely varying quantities of granulocytes and numbers of transfusions, and (*d*) included patients in whom bacterial infection had not been documented. However, the majority of studies concluded that granulocyte transfusions appeared to be beneficial, and we wished to evaluate this modality in pediatric patients. In the discussion of this study with the Biostatistics Division of our hospital, the clinical question was: how can we best evaluate the efficacy of granulocyte transfusions in neutropenic children with bacterial infections? We wished to have a controlled study, but one in which, if transfusions were beneficial, the answer would be available as soon as possible. We were asked to provide the estimated number of patients available for such a study and the clinical background and limitations.

## Relationship of Neutropenia to Infections in Leukemia

Infections secondary to neutropenia remain a major cause of death in cancer patients (1). Both the duration

[1]Supported by Public Health Service grants RR-05584 (Division of Research Resources) and CA-21765 (National Cancer Institute) from the National Institutes of Health, Department of Health, Education, and Welfare, and by ALSAC.

[2]Department of Hematology and Oncology, St. Jude Children's Research Hospital, Memphis, TN.

[3]I wish to thank the following people for their contributions to the design of the protocol and their assistance in gathering data: *Drs. Ronald G. Strauss* (University of Iowa, Iowa City), *Sandor Feldman* (St. Jude Children's Research Hospital), and *Michael Ossi* (Auburn Medical Association, Inc, MA).

**Reprint requests to:* Dr. Elizabeth I. Thompson, St. Jude Children's Research Hospital, 332 N Lauderdale, PO Box 318, Memphis, TN 38101.

TABLE 1.—Distribution of patients with bacteremia

| | No. of patients with bacteremia | | |
|---|---|---|---|
| | Induction therapy | Initial relapse | Totals |
| Acute lymphocytic leukemia | 6 | 8 | 14 |
| Acute myelocytic leukemia | 2 | 2 | 4 |
| Non-Hodgkin's lymphoma | 2 | 1 | 3 |
| Totals | 10 | 11 | 21 |

and severity of neutropenia contribute to the likelihood of infection (2). Gram-negative bacilli, *Staphylococcus aureus*, and *Candida albicans* are the most frequent causative organisms.

The highest mortality rates from infection occur in patients with persistent neutropenia. In one study of cancer patients treated with gentamicin for gram-negative bacillary infections, response to therapy was significantly different depending on granulocyte counts. Only five of 22 (23%) infections responded favorably in patients with granulocyte counts remaining < 1000 neutrophils/μl; 19 of 24 (79%) infections responded favorably in patients with granulocyte counts > 1000 neutrophils/μl (3).

In large series of neutropenic leukemic patients with infections mortality rates are ≥ 50%. However, these series include patients with progressive leukemia and were reported before recent improvements in antibiotic therapy. Therefore, in conjunction with the Department of Infectious Disease, we reviewed our recent experience at St. Jude Children's Research Hospital, including only patients who had a high likelihood of achieving remission and who also had documented infection and neutropenia. We selected for review patients who would have been candidates for granulocyte transfusions in our proposed study: children with leukemia (or lymphoma with leukemic conversion) who had severe neutropenia and bacteremia and who were in induction therapy, first relapse, or remission. By restricting eligibility for this study to patients with a high probability of achieving remission, we hoped to eliminate as much as possible the influence of progressive leukemia on the potential recovery from infection.

From May 1974 to May 1975, 367 children with acute leukemia and lymphoma with leukemic conversion were treated at our hospital. There were 68 episodes of bacteremia among these patients. However, limiting episodes of bacteremia to those patients who had severe neutropenia, 21 episodes occurred among 19 patients (absolute granulocyte count ≤ 500/μl in 18 and 592/μl in one). Table 1 shows the distribution of these patients.

Table 2 shows the organisms isolated in the 21 episodes of bacteremia in these 19 patients. Five patients died during the initial episode of bacteremia; all five remained neutropenic and died of infection within 7 days. A sixth death occurred in one of the two patients who had a second episode of bacteremia. No deaths occurred among patients with the nine episodes of *S. aureus* infection. Thus, the overall mortality rate from these 21 episodes was 28.5%, but if only gram-negative bacilli or multiple organisms are considered, the mortality rate was 42%.

TABLE 2.—Causative organisms and outcome of infection in 367 children with acute leukemia and lymphoma with leukemic conversion

| | No. of— | |
|---|---|---|
| | Episodes of infection | Death |
| *S. aureus* | 9 | 0 |
| *Pseudomonas aeruginosa* | 6 | 2 |
| *Escherichia coli* | 3 | 1 |
| Herella sp. | 1 | 1 |
| *S. aureus* + *P. mirabilis* | 1 | 1 |
| *E. coli* + enterococci | 1 | 1 |
| Totals | 21 | 6 (28.5%) |

The median duration of neutropenia was 13 days (range, 4–32 days) for the 15 patients who survived. Serial blood cell cultures were available in 12 patients. Three of these 12 patients had persistent bacteremia after > 48 hours of appropriate antibiotic therapy. The median duration of antibiotic therapy was 12 days (range, 6–20 days). Fever persisted for a median of 6 days (range, 2–29 days). The median length of hospitalization was 13 days (range, 7–48 days). Thus, morbidity as well as mortality was significant in these children, and recovery of bone marrow function was delayed.

## Previous Use of Granulocyte Transfusions in Patients With Neutropenia

Space does not permit a review of all previous studies of granulocyte transfusions reported by 1975; the three following studies were chosen as examples.

Graw et al (4) infused granulocytes harvested by either continuous-flow centrifugation or filtration leukopheresis to granulocytopenic patients with culture-proven gram-negative septicemia. Their controls were patients for whom a donor was not available. Eighteen of 39 (46%)

granulocyte-treated patients and eight of 27 (30%) control patients survived; the difference was not statistically significant. However, 12 patients who received four or more transfusions survived, whereas in a comparable group of controls, only five of 19 survived (100% versus 26%). When only those patients in whom granulocyte counts did not return to normal were evaluated, one of 11 control (9%) and five of 12 treated (42%) patients recovered. This difference is significant, and Graw et al (4) concluded that repeated granulocyte transfusions improved the management of severe infection in patients with persistent granulocytopenia.

Lowenthal et al (5) treated 41 granulocytopenic patients with cells collected by continuous-flow centrifugation. Bacterial infection was assumed but not always documented. There was wide variation in the number of cells per transfusion. In 40% of patients, fever decreased coincident with transfusions. These patients received nearly four times the number of cells as those who did not respond.

In a controlled study, Fortuny et al (6) gave $6 \times 10^9$ granulocytes/transfusion/day for 1–8 days. They noted no advantage in the transfused versus the antibiotic group with regard to either mortality or resolution of infection. Seventy-six percent of the patients treated with antibiotics alone survived, a higher proportion than in most series.

In summary, previous studies of granulocyte transfusions were encouraging, but necessarily raised many questions. Should all neutropenic patients with sepsis receive transfusions or only those with persistently positive cultures? What dose of granulocytes and frequency are minimal? Would better monitoring of antibiotic therapy decrease the need for transfusions? Since children often have a shorter duration of neutropenia, are indications for granulocyte transfusions the same as for adults? Would donor-recipient histologic antigen matching improve survival of transfused granulocytes? Do immunosuppressed patients form antibodies to infused granulocytes?

## CLINICAL PLAN OF THE PROTOCOL

Obviously, all of the many questions could not be addressed in a single study. With carefully selected patients, a trial seemed reasonable with all patients receiving similar supportive care and randomization to receive or not receive at least four consecutive daily granulocyte transfusions within 24 hours of diagnosis of bacterial infection. Severe neutropenia would be documented, and because of difficulties in interpreting the significance of pulmonary infiltrates and bacteriuria, the study would be restricted to patients with positive blood or cerebrospinal fluid cultures. One difficult problem is ascertaining whether all patients are receiving "optimal" antibiotic therapy. Although the kinetics of antibiotic serum levels (7) and the correlation of in vitro minimal inhibitory concentration of antibiotics for specific organisms are still being defined, we elected to measure these levels in all patients in the study. We also elected to harvest granulocytes by a centrifugation method (Haemonetics-30 cell separator) and use family member donors, with histologic antigen typing correlated retrospectively.

The proposed clinical plan, then, was as follows. To be eligible, a patient would have to have $< 500$ neutrophils/$\mu$l and a positive blood or cerebrospinal fluid culture within 24 hours of admission into the hospital with suspected sepsis. Antibiotic therapy (oxacillin, gentamicin, and carbenicillin) would be given immediately to all children. Eligible patients would be restricted to those with acute lymphocytic leukemia, acute myelocytic leukemia, or lymphoma with leukemic conversion, who were in induction therapy, first relapse, or remission. Informed consent would be obtained from the patient and/or parents and from potential family member donors who were ABO cross-match-compatible with the patient. Patients would then be randomized to receive (treatment 1) or not receive (treatment 2) four consecutive daily granulocyte transfusions, each containing at least $1 \times 10^{10}$ granulocytes/m$^2$. The cells would be obtained by a Haemonetics-30 cell separator and the donor would receive hydroxyethyl starch and corticosteroids. Levels and minimal inhibitory concentrations of antibiotics for the patient's specific organisms would be monitored in both groups. The cells would be irradiated before transfusion.

## STATISTICAL CONSIDERATIONS

When this clinical plan was proposed, we emphasized in our discussions with the statisticians that a primary concern of the study would be to ascertain as soon as possible whether the use of granulocyte transfusions reduced the mortality rate. In essence, we were stating that we wished to frequently review the data as the study progressed. The concept in fixed-sample analysis is to derive beforehand the number of patient randomizations which will give statistically significant results. Early analysis of data in a fixed-sample trial therefore violates the statistical principles of the design.

An alternate method, and the one selected for this trial, is a standard form of sequential design: the "repeated significance test," a special type of restricted sequential design (8). The plan is restricted in that an upper limit is placed on the maximum number of patients required. In this plan, mortality results are under constant and close supervision.

To select sequential design as the statistical tool for a clinical trial, several limiting criteria must be met. First, the effect of therapies being tested should be easily defined. Since our main concern was death versus survival,

the protocol met this limitation. Second, the results of therapy should be measurable within a short time interval, which would also be true for our study. Generally, only one randomization is evaluated, but patients may be stratified. Since there was no incidence of mortality in the review of our patients with *S. aureus* infections, we wished to evaluate this subgroup separately. This could be done with sequential design as follows.

## Statistical Study Design

Patients entering the trial would be randomized and paired in two different strata: one stratum for patients with gram-negative or multiple-organism infections and another stratum for all other patients. The therapy would be designated: *treatment 1*—standard antibiotic therapy, and *treatment 2*—standard antibiotic therapy plus granulocyte transfusions. The randomization would be balanced within each stratum but, to avoid potential bias, the nature of the balance would not be divulged to the clinicians. Thus, our Biostatistics Division would be notified that an eligible patient had gram-negative or positive bacteremia, and the clinician would receive notification for treatment 1 or 2.

The trial would proceed by randomly pairing patients receiving the two different treatments in the order that they were entered in the trial, and by plotting the outcome on a chart similar to that shown in figure 1. The boundaries shown are calculated according to statistical significance in sequential design, ie, if one of the boundaries is reached, significance may be declared. Regardless of stratum, each pair of patients that resulted in a "preference" for one of the two treatments (survival in one treatment and death in the other) would be plotted on the figure. Starting at the origin 0, for each preference for treatment 1, the line would progress 1 unit to the right and 1 unit upward; for each preference for treatment 2, the line would progress 1 unit to the right and 1 unit downward. The trial would continue as long as none of the boundaries was reached. Because of the declared boundaries, this trial would have the following characteristics: (*a*) a decision would be potentially possible after no more than seven preferences were recorded; (*b*) the maximum number of preferences would be estimated at 25; and (*c*) the expected number of preferences would be estimated at 13.

Sequential trial design was chosen to avoid continuing the trial beyond that point at which one therapy would be determined to be significantly better than the other in terms of mortality rates. It is important to recognize, however, that if treatments 1 and 2 do not markedly differ, this kind of study may take longer and require *more* patients than fixed-sample analysis. Also, even if the two mortality rates are very different, the probability of preference may not be much different from 0.50. For

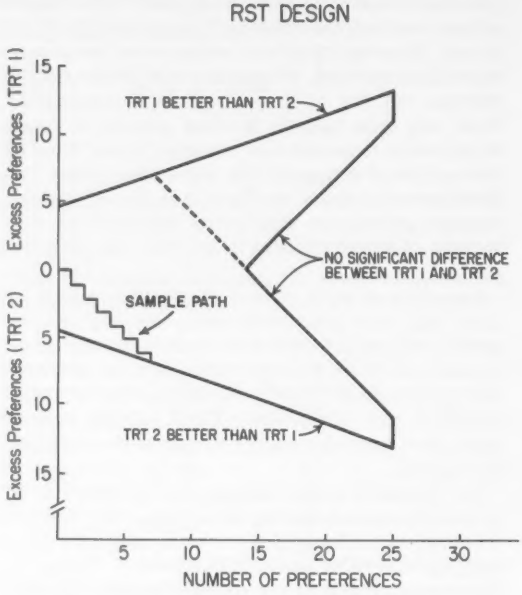



FIGURE 1.—The repeated significance test (RST) design with a sample path. The path would be plotted if treatment 2 (TRT 2) were better than treatment 1 (TRT 1) for the 1st 7 pairs that were evaluated. The path crosses the line. If this event occurred, treatment 2 could be declared significantly superior to treatment 1.

example, if the mortality rates for treatments 1 and 2 are 0.55 and 0.20, respectively, the probability for a preference of treatment 1 or 2 for any random pair is still only about 0.53. That is, about one half of the pairs will result in "ties" (both fatal or both nonfatal outcomes). Therefore, if N pairs are entered, only N/2 will provide information on differences. On this basis, one can estimate the minimum, maximum, and expected numbers of *pairs* required under these conditions as 14, 50, and 26 respectively, and the number of patients required as 28, 100, and 52, respectively. If, for example, 25 patients/year are entered in the study, the trial could have a minimum duration of 1 year, a maximum duration of 4 years, and an expected duration of 2 years.

Soon after this trial was proposed, studies were published (9,10) suggesting that, for patients with a 10- to 14-day duration of neutropenia, granulocyte transfusions produced no measurable difference in mortality. In addition, potentially eligible patients who were observed while the protocol was being planned survived without granulocyte transfusions, except for one patient who died within 4 hours of admission. Therefore, patient selection in this protocol will be modified to evaluate patients likely to have a duration of neutropenia of > 10–14 days.

In conclusion, sequential design is a potential alternative to fixed-sample analysis in clinical trials which fit the limitations of this statistical approach. It may be particularly advantageous in trials of new forms of therapy where there are ethical concerns for determining efficacy of potentially beneficial therapy as rapidly as possible. In utilizing this statistical approach, however, the risk must be accepted that, if two forms of therapy are similar in efficacy, the trial may require more patients and a longer duration than are required with fixed-sample analysis. This risk may be reasonable when the primary concern is to determine whether one form of therapy is markedly superior to another.

## REFERENCES


1. Hersch EM, Bodey GP, Nies BA, et al. Causes of death in acute leukemia. JAMA 193:105–109, 1965.
2. Bodey GP, Buckley M, Sathe YS, et al. Quantitative relationships between circulating leukocytes and infection in patients with acute leukemia. Ann Intern Med 64:328–340, 1966.
3. Bodey GP, Middleman E, Umsawadi T, et al. Infection in cancer patients: results with gentamicin sulfate therapy. Cancer 29:1697–1701, 1972.
4. Graw RG, Jr, Herzig G, Perry S, et al. Normal granulocyte transfusion therapy: treatment of septicemia due to gram-negative bacteria. N Engl J Med 287:367–371, 1972.
5. Lowenthal RM, Grossman L, Goldman JM, et al. Granulocyte transfusions in treatment of infections in patients with acute leukemia and aplastic anaemia. Lancet 1(7903):353–358, 1975.
6. Fortuny IE, Bloomfield CD, Hadlock DC, et al. Granulocyte transfusion: a controlled study in patients with acute nonlymphocytic leukemia. Transfusion 15:548–558, 1975.
7. Evans WE, Feldman SF, Barker LF, et al. Use of gentamicin serum levels to individualize therapy in children. J Pediatr 93:133–137, 1978.
8. Armitage P. Sequential medical trials, 2nd ed. New York, NY, Halsted Press, John Wiley and Sons, 1975.
9. Herzig RH, Herzig GP, Graw RG, Jr, et al. Successful granulocyte transfusion therapy for gram-negative septicemia. A prospectively randomized controlled study. N Engl J Med 296:701–705, 1977.
10. Alavi JB, Root RK, Djerassi I, et al. A randomized clinical trial of granulocyte transfusions for infection in acute leukemia. N Engl J Med 296:706–711, 1977.

# Patient Heterogeneity in Clinical Trials

**Richard Simon** [1,*]

## SUMMARY

**Interpretation of therapeutic results is complicated by variability in response among patients. This paper reviews fundamental statistical principles for the design of clinical trials. These methods seek to evaluate relative therapeutic efficacy in the presence of patient heterogeneity. Statistical science has more to offer therapeutics than significance tests among "comparable" treatment groups. The role of randomization and stratification is reviewed. The importance of study design, including patient eligibility and therapeutic standardization, to the generalization of conclusions is discussed.**

[Cancer Treat Rep 64:405–410, 1980]

Clinical trials are characterized by heterogeneity in patients treated and by variability in responses obtained. Such heterogeneity has caused statistics to assume considerable importance in the conduct of modern clinical trials. In fields in which response is completely predictable and unvarying, there is no need for statistical design or analysis. Clinicians, pathologists, and statisticians all have developed methods for dealing with heterogeneity. Many such methods are based on the development of patient categories and the comparison of therapeutic results within such categories or among mixed groups "comparable" with regard to the category variables. The categories are generally based on site of primary tumor, stage, and histologic classification. This paper reviews some of the statistical methods by which we can deal with patient heterogeneity.

## Comparability

The notion of "comparability" is basic to the evaluation of competing therapies. It is meaningless to compare treatments administered to patient populations that differ with regard to important prognostic factors. For most clinical studies, however, comparability denotes equal heterogeneity, although it is impossible to achieve complete comparability because of the number of factors that may affect response, many of which are unknown or unmeasurable by practical means. Comparability itself is not a statistical concept. A fundamental statistical principle for dealing with such heterogeneity is to define the treatment groups so that the uncontrollable sources of variability of the estimator of treatment difference have a symmetric and known distribution. This point is sometimes inadequately appreciated. It is frequently stated that the purpose of randomization is to ensure that treatment groups have comparable prognostic factors. With known and measured prognostic factors, this objective can be accomplished at least as well by matching. With unknown or unmeasured prognostic factors, randomization does not ensure comparability, but it does ensure a symmetric and known distribution of error which is used to evaluate the reliability of the estimator of treatment difference. Without randomization, uncontrolled prognostic factors may produce a systematic bias that cannot be estimated. Consequently, the analysis of nonrandomized studies should include an exhaustive evaluation of possible sources of noncomparability. In a recent paper, Rubin (1) pointed out that even from the Bayesian point of view, the analysis of causal effects is problematic in the absence of randomization.

## STRATIFICATION

Stratification is widely practiced in randomized clinical trials. This procedure usually partitions patients into mutually exclusive subsets based upon pretreatment characteristics thought to affect prognosis. For example, in a comparison of treatments A and B for metastatic

[1]Biometric Research Branch, Clinical Oncology Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD.

**Reprint requests to:* Richard Simon, PhD, Biometric Research Branch, Clinical Oncology Program, Division of Cancer Treatment, National Cancer Institute, Bldg 10, Rm 3B16, National Institutes of Health, Bethesda, MD 20205.

breast cancer, the appropriate factors may be dominant metastatic site, age, and performance status (fig 1). Each subset or stratum is then subjected to a pseudorandomization such as the permuted block design. A permuted block of length 2k is a sequence of As and Bs containing exactly kAs and kBs. The 20 possible permuted blocks of length 6 are shown in figure 2. The treatment assignment is determined by a succession of randomly selected permuted blocks within each stratum. If there are three treatments, permuted blocks of k instances of each of three letters are used. Generally, the permuted blocks of distinct strata are prepared independently of each other. A pseudorandomization mechanism such as permuted blocks is necessary because pure randomization within each stratum is equivalent to no stratification at all.

Stratification is seen as ensuring a greater degree of comparability of the treatment groups than would be obtained by pure randomization. It accomplishes this objective; however, unless the important stratification factors are incorporated into the analysis, the full potential of stratification for increasing the sensitivity of the trial is not realized. It is general practice to stratify and then to ignore the stratification factors in the analysis. This practice results in a more precise estimator of treatment differences than if stratification were not employed. The precision of the estimator is underestimated, however, because the known sources of variability embodied in the stratification factors are considered unknown sources of noise.

The oversimplified illustration in figure 3, which assumes two strata and two treatments, may help to clarify this concept. The upper graph depicts the frequency function (histogram) of survival for each of the two treatment groups within each of the two strata. For purposes of illustration, the concept of censored survivals is ignored. It is clear that stratum II patients have a better prognosis than stratum I patients, because their frequency functions are displaced to the right. Also, within each stratum treatment B seems more effective than treatment A. The desirable analysis consists of calculating an estimate $\Delta_I$ of the treatment difference within stratum I and an estimate $\Delta_{II}$ of the treatment difference within stratum II. A weighted average of $\Delta_I$ and $\Delta_{II}$ is used as the overall test statistic for evaluating the statistical significance of the treatment difference. Significance tests could be performed within each stratum, but the power of these tests usually will be small because of limited numbers of patients. The variability associated with the weighted average of $\Delta_I$ and $\Delta_{II}$ is estimated from the observed within-stratum within-treatment group variability. Stated differently, the variability associated with the test statistic is determined from the estimated variability for each of the four curves at the top of figure 3. The treatment differences in the case shown are reasonably large relative to the variability of the individual curves, so if the number of patients is reasonable, a small significance level will be obtained.

In the bottom of figure 3, the stratum distinctions are ignored. Each of the two curves represents a frequency function of survival for a pooled treatment group consisting of patients from both strata. Although the two treatment groups may be perfectly comparable with regard to the stratification variable, if the pooled groups are compared directly, a very powerful significance test will result.

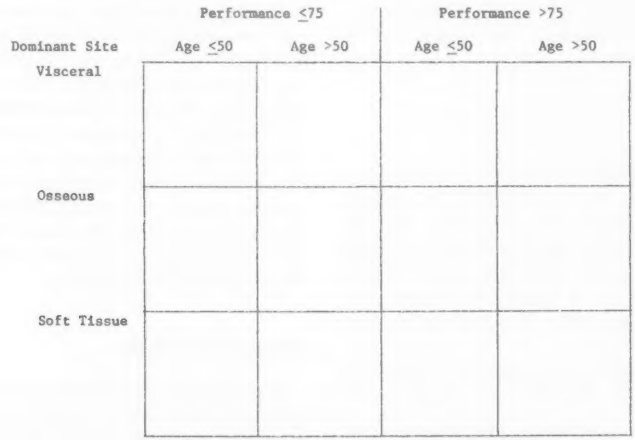



FIGURE 1.—Stratification design for metastatic breast cancer.

| | |
|---|---|
| AAABBB | BBBAAA |
| AABABB | BBABAA |
| AABBAB | BBAABA |
| AABBBA | BBAAAB |
| ABAABB | BABBAA |
| ABABAB | BABABA |
| ABABBA | BABAAB |
| ABAAAB | BAABBA |
| ABBABA | BAABAB |
| ABBBAA | BAAABB |

FIGURE 2.—All permuted blocks of length 6 for 2 treatments.

Although limited stratification is desirable, overstratification can be detrimental to a trial. Numerous strata may contain only a single patient at the conclusion of the trial. Balance of the most important factors may be impaired by the inclusion of secondary factors in the design. Even the total number of patients assigned each treatment may be very unbalanced. In such trials, overstratification becomes equivalent to no stratification (2). These detrimental results occur with overstratification, because many strata do not contain enough patients to fill the first permuted block.

Since the blocked randomizations for different strata are prepared independently, gross imbalances may result. In figure 4 the sequence of treatment assignments is generated for the stratification design (fig 1) by independently and randomly selecting permuted blocks of length 6 for two treatments: A and B. The check marks indicate the treatment assignments and distribution of stratification factors for the first 26 patients. Although the sizes of the treatment groups are well-balanced in this case (14 As to 12 Bs), the distributions of dominant tumor site and performance status are quite unbalanced, as shown at the bottom of figure 4, because the study contained too few patients for the number of strata and block length used. In this case perhaps the problem could have been avoided by using blocks of 2 or 4. In general, however, even if blocks of length 2 are used, the same

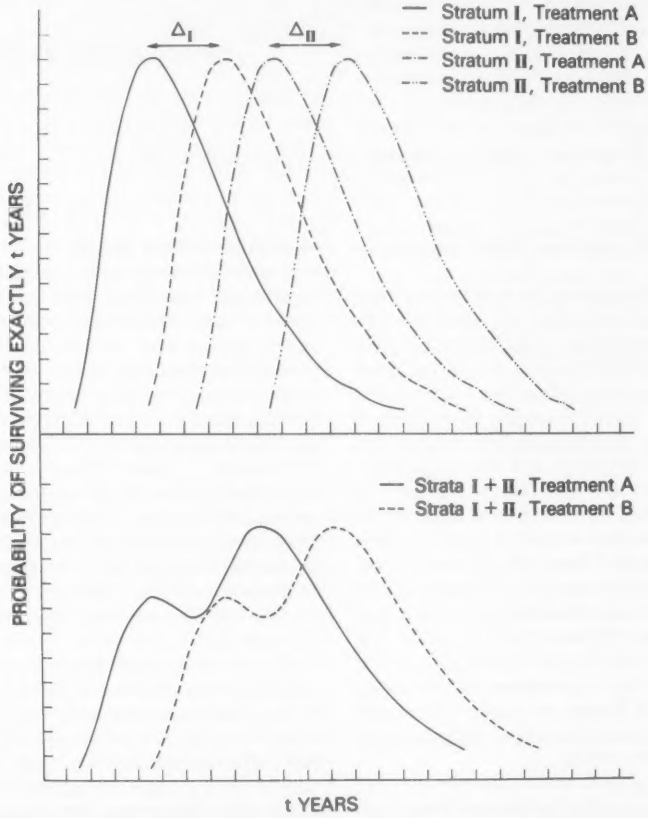



FIGURE 3.—Use of stratification factors in analysis.

| Dominant Site | Performance ≤75 Age ≤50 | Performance ≤75 Age >50 | Performance >75 Age ≤50 | Performance >75 Age >50 |
|---|---|---|---|---|
| Visceral | A✓ A✓ B A B B | A✓ A✓ A✓ B B B | A✓ A✓ A✓ B A B | B✓ A✓ A✓ B B A |
| Osseous | A✓ A✓ B B B A | B B A A A B | B✓ B✓ A A B A | A✓ B✓ B✓ A A A |
| Soft Tissue | B✓ B A A B A | B A A B B A | A✓ B✓ B✓ B✓ A A | B✓ B✓ A A A B |

| | | A | B |
|---|---|---|---|
| Dominant Site | V | 10 | 1 |
| | O | 3 | 5 |
| | S | 1 | 6 |
| Performance | ≤75 | 7 | 1 |
| | >75 | 7 | 11 |
| Age | ≤50 | 8 | 6 |
| | >50 | 6 | 6 |
| | | 14 | 12 |

FIGURE 4.—Imbalances in an overstratified study.

problem may result if the degree of stratification is too great.

Relatively recent developments in stratification have been reported by Pocock and Simon (2), Sedransk (3), Harville (4), Taves (5), and Begg and Iglewicz (6). Peto et al (7) have dismissed stratification for all but small studies, and view it as a complication that has been rendered unnecessary by the introduction of methods of analysis that adjust for prognostic factors. An editorial in the British Medical Journal (8) put this viewpoint in these terms: "Nowadays it is generally unnecessary to randomize patients within separate groups according to prognostic factors, since these are better allowed for retrospectively in the analysis." One such adjustment procedure, previously described, consists of determining an estimator of within-stratum treatment difference, and using a weighted average of these estimators as the test statistic. The commonly used Mantel-Haenszel test (9) is a specific example of such a procedure for comparing response rates, survival curves, or remission-duration curves. Such adjustment procedures exemplify a broader class of regression models (10,11).

Peto et al (7) point out that adjustment methods can be used whether or not stratification was performed during treatment assignment. They grant that stratification does help to ensure that the stratification factors will be well distributed among the treatment groups so that a significance test will be more powerful, but they argue that for large clinical trials the expected power will be nearly as good even without stratification.

Not all medical statisticians share this viewpoint. For small clinical trials some degree of stratification by important variables seems desirable. Even for larger studies, the expected power of significance tests is not the only notion of power to be considered. Viewed in one way, stratification is an insurance policy against low probability imbalances that can ruin or severely decrease the sensitivity of a study. Any particular study will be performed only once, and if the treatment effect is nearly confounded with the effect of a prognostic factor, it will be little consolation to know that the expected power was adequate (12).

The use of prospective stratification tends to avoid situations in which the medical community is not convinced by the conclusions of a study because it may not understand the analysis used to adjust for a lack of comparability. If stratification is employed, the method of analysis may be the same and still may not be widely understood, but it is less likely that a troublesome lack of comparability will exist to arouse skepticism.

Finally, it should be noted that the validity of an adjustment technique depends upon the adequacy of the assumed statistical model. The number of factors that can be incorporated may be severely limited by the number of patients treated. There may be great ambiguity concerning which variables to adjust for. Adjustment for all variables may entail considerable loss of power, and selection of adjusting variables by stepwise procedures or informal methods makes the resulting conditional significance level arbitrary. Using limited stratification, the variables and their representations in the analysis can be uniquely specified by the design. The resulting significance level can also be interpreted on the basis of replications of the randomization for the same patients.

Some investigators give undue and detrimental emphasis to extensive stratification. Large definitive studies identifying the variables which exert important independent contributions to prognosis are very valuable. Such studies often demonstrate that potential stratification variables are so related that only a few need be included for the purposes of design and analysis.

## GENERALIZATION OF CONCLUSIONS

If the conclusions of a clinical trial were not generalizable to patients other than those actually entered in the study, clinical research would be futile. The only statistical basis for generalization is the assumption that the patients studied constitute a random sample from a larger population. Such random selection of patients is rarely if ever practiced, but with certain stipulations it often is reasonable to generalize our conclusions as if our patients were so selected. For several aspects of study design to be described, it is important to bear in mind the issue of generalization of conclusions.

The determination of patient eligibility requirements is a crucial part of study design. Such requirements are usually based on two assumptions: (*a*) the patients are able to tolerate any of the treatments and (*b*) potential benefits outweigh potential risks. Often such considerations do not define specific eligibility requirements, and questions arise about whether it is desirable to include and stratify for various small distinguishable subgroups of patients. An important implication of including such subgroups should be recognized. Because of limited sample sizes, it will rarely be possible to compare treatment efficacy for such subgroups separately, and the basic conclusion of the study will result from the overall comparison. Although the overall comparison may be adjusted for stratification factors, the conclusion based on this comparison can be generalized only to a mixture of the same kinds of patients studied in the trial. This holds true unless one is willing to make the generally untenable and inadequately supported assumption that relative treatment efficacy is the same for all subgroups. Consequently, it is important that the patient eligibility requirements define a medically meaningful group of patients, so that in reporting the results of the study a coherent statement can be made as to whom the conclusions apply. Hill (13) emphasized this point in his description of a Medical Research Council trial of streptomycin: "In short, the questions asked of the trial were deliberately limited and these 'closely defined features were considered indispensable, for it was realized that no two patients have an identical form of the disease and it was desired to eliminate as many of the obvious variations as possible.' This planning . . . is a fundamental feature of the successful trial. To start out upon a trial with all and sundry included, and with the hope that the results can be sorted out statistically in the end is to court disaster."

The same considerations apply to two other important issues of study design. The first issue is standardization of surgery and radiotherapy in chemotherapy adjuvant studies. It is generally agreed that if we are studying the efficacy of chemotherapy, the administration of chemotherapy should be standardized. Standardization of the surgery or radiotherapy to which the chemotherapy is an adjuvant is often more difficult. However, if the local therapies are not standardized, the conclusions apply, if at all, only to a mixture of patients receiving the same types of local therapy as those studied. The conclusions may not be valid for any specific patient in a mixed population for the following reason. Chemotherapy may benefit patients receiving extensive surgery with limited radiotherapy, but may be detrimental to patients receiving limited surgery with extensive radiotherapy. Limited numbers of patients may not permit reliable evaluation of the effect of chemotherapy within the subgroups of patients receiving similar local treatment. The variability introduced by nonstandardization makes the conclusion that chemotherapy has no significant positive or negative effect erroneous for all patients. Standardization should not be carried so far, however, as to make it impossible to conduct a reasonable study with adequate numbers of patients.

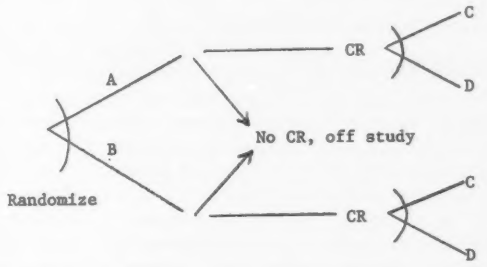



FIGURE 5.—A 2-stage clinical trial with 4 treatment groups.

The final important issue in the context of generalization of conclusions is that of patient disqualifications. Because we wish to evaluate treatment efficacy, we are tempted to disqualify patients who are not adequately treated for one reason or another. This is a dangerous practice because of the potential biases it permits. In most situations, it is unlikely that the patients who do not complete treatment can be viewed realistically as a random selection of all patients assigned that treatment. Generally, their prognoses are poor. The treatments being compared may differ in the kinds and quantities of potential disqualifications they produce; therefore disqualifications based on anything but eligibility criteria are likely to introduce bias into the evaluation. The only fully reliable approach is to include all patients actually eligible for the study in the analysis. One is then evaluating the therapeutic *strategy* of attempting to administer treatment A versus the strategy of attempting to administer treatment B. With this approach, conclusions may be more clinically relevant and may be generalized to broader patient populations.

## REFERENCES


1. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat 6:34–58, 1978.
2. Pocock SJ, and Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 31:103–115, 1975.
3. Sedransk N. Allocation of sequentially available units to treatment groups. *In* Proceedings of the 39th International Statistical Institute, 1973, vol 2, pp 393–400.
4. Harville DA. Nearly optimal allocation of experimental units using observed covariate values. Technometrics 16:589–599, 1974.
5. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. Clin Pharmacol Ther 15:443–453, 1974.
6. Begg CB, and Iglewicz B. A treatment allocation procedure for sequential clinical trials. Biometrics 36:81–90, 1980.
7. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 34:585–612, 1976.
8. Editorial. Randomized clinical trials. Br Med J 4:1238–1239, 1977.
9. Mantel M, and Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. JNCI 22:719–748, 1959.
10. Cox DR. Regression models and life tables. J Stat Soc Br 34:187–220, 1972.
11. Cox DR. Analysis of Binary Data. London, England, Methuen, 1970.
12. Lasagna L. Randomized clinical trials. N Engl J Med 295:1086–1087, 1976.
13. Hill AB. The clinical trial. Br Med Bull 7:278–282, 1951.

## DISCUSSION I: Basic Design Considerations *

**Dr. Elashoff:** In his paper, Dr. Gehan described his favorite ways of adjusting for prognostic factors in a historical controlled framework. He outlined a comparison of FAC/BCG against a surgery and a radiotherapy historical study intended to assess the effectiveness of the adjuvant chemotherapy. There could have been another selection factor, namely that chemoimmunotherapy was imposing or would impose considerable stress, and the study design assumed that a patient ought to be able to stand up under an aggressive chemoimmunotherapy routine. Such stress is a difficult element to bring into the analysis per se. With respect to comparability or the choice of prognostic factors, an index of prognosis was obtained from a study of the control group, and these variables were then used to analyze for treatment effects. They were brought into the regression equation which was applied to identify treatment effects, apparently without looking at a possible treatment group by prognostic variable interaction, or by whether the prognostic variables were indeed going to be useful for the FAC-BCG group. In any event, it is difficult to adjust for factors by regression equation; it is hard enough in classical multiple regression. I never seem to have too much success. Others undoubtedly have a different experience.

Another approach to the use of prognostic factors is to determine mortality rates for the various levels of each prognostic factor. If there seems to be variability at the different levels, that variability might be used in a regression equation or in another approach. An effort should be made to cluster these prognostic variables in a way that differs from Dr. Gehan's. Everyone has a favorite way of searching for prognostic factors, and as long as one is sensible about what one does and knows what things can go wrong, there is probably no harm in any method. There is no best way.

When Dr. Gehan was indicating on one of his tables that the treatment variable was the strongest of all the variables that entered into the regression equation, he did not make it clear whether the treatment was adjusted for covariance.

Dr. George, in his overview of sequential clinical trials, has given his definition of "sequential," namely, any analysis of accumulating data. Again I was looking for a more detailed description of the followup sequential clinical trials. We had an "event" sequential followup trial, and at times a sequential followup trial, and I would have liked some discussion of alternative ways of conducting a sensible followup clinical trial for various cancers.

Dr. Simon talked about a number of important issues—patient heterogeneity, and what randomization will be achieved—and gave us some ideas for random allocation. He also examined the question of prerandomization stratification, which embraces many different views. He recommended a little stratification and recommended "floating" with randomization and adjusting by one or another of the regression models.

The paper by Dr. Zelen (presented but not published here) was a novel, innovative, and welcome challenge. Unfortunately, I got the paper too late to discuss it with Dr. Jeremy Thompson, who is head of our Committee on Human Experimentation at UCLA, to get his reaction to it. It would be most interesting to learn how our Committees on Human Experimentation would respond to the ideas conveyed in Dr. Zelen's talk. In any case, I believe these designs merit close examination.

**Dr. Ellison:** I am a statistically unsophisticated clinician and, as such, wish to raise questions rather than comment on the validity of the ideas presented. I am curious about the analysis of prognostic factors. In a way, we may fall into the trap of the self-fulfilling prophecy, particularly if we look at a large group of patients and say, "These are the ones who did better; therefore, these are the ones who will do better in this study." As we go from study to study we have to be very careful that the models we decide upon, or those the statisticians derive by one technique or another from the clinical data, remain valid from study to study and ensure that the variations we see arise from variations in the patients rather than variations in the way we are treating them or the methods of assessing the effects. Should we not look at things in reverse? Instead of saying "This parameter becomes the most important one and this is the second most important, etc," should we not look at the patients who did best and at the patients who lived longest and so derive the common characteristics of these patients?

We have been told that too many strata are bad because we end up with an unrandomized study. How many are too many? Is this a constant figure? Is it related to the number of questions we are asking in a study, the number of "arms" into which we are randomizing patients? Also, we should recognize how much patient-risk factors are complicated by interaction—not just interaction of the risk factors, but interaction with the various elements of many complex studies. Experience in protocols of acute leukemia has taught me that rarely are we just giving

*Reprint requests to: Dr. Robert Elashoff, Department of Biomathematics, University of California, Los Angeles, CA 90024.

a drug or a combination of drugs and looking at the remission rate or the duration of survival. We are not asking single questions, we are asking multiple questions concerning drugs, drug dose, drug schedules, and the use of maintenance and intensification treatment. We also have difficulty in our ability to analyze for factors operative at different times in the study and in the patient's course. Thus most of the prognostic features that have been obtained, whether from logistic regression or from other techniques, are those that were present before the initial treatment. What about the influence of factors that arise subsequently? Which results derive chiefly from preinduction characteristics, whether they are biological, biochemical, or physical?

What unknowns do we introduce during induction? We are giving known drugs but are indulging in polypharmacy; these same patients are also receiving antibiotics, analgesics, narcotics, antiemetics, and a number of other drugs that affect the metabolism of the cytotoxic drugs. In addition to this polypharmacy, there are changes in the patient's immune status and hematopoiesis due to the use of cytotoxic drugs. Thus there are complicated questions made more confusing by the complicated factors impinging upon these complicated questions.

To what degree is the survival of patients or the remission duration related to how we handle "intensification" treatment methods or how we handle maintenance treatment? What part of the survival is related to the status of the patients at the time that we start maintenance? Dr. Miller mentioned that those children with ALL who had fewer blasts at Day 14 after the initiation of induction therapy did better. Should we compare marrow involvement at a fixed time (such as this)? Should we look also at the length of time to the "elimination" of obvious blast cells, at the degree of elimination of leukemic cells, and at the duration of suppression of normal marrow?

How can clinicians work most efficiently and most effectively with statisticians, not only in planning studies but also in carrying them out and in assisting with the analysis by providing the appropriate data? I thought of this with regard to the sequential clinical trials design. It seems to me that if we terminate studies as early as suggested by this design, we will collect data that other clinicians will not accept as meaningful, although there is statistical significance. What educational process will be required to achieve acceptance for these conclusions? Perhaps my insecurity resulting from not knowing the stopping point in advance makes it difficult for me to follow the design and see how to achieve the appropriate number and type of entries into the study. In some studies we have difficulty in persuading other physicians to allow their patients to be entered, and further trouble persuading patients to enter. If we do not know how long the study will be, will this be more difficult? Do sequential clinical trials give us an answer when we are asking multiple and/or complicated and/or interrelated questions? Do they give us primarily a "yes-no" answer, or can they be used for all of these questions?

Patient disqualification is a problem that we discuss frequently as clinicians. If we say that the only disqualification is for a patient who is ineligible, what do we do when there has been gross disregard of protocol stipulations? If this major protocol violation results from drug toxicity or, as is frequently seen, from the patient's refusal to be made sick, this can be classed as a failure of the treatment. What do we do in the event of gross disregard of treatment because of accidents or because someone has not read the protocol? Isn't this grounds for ineligibility? Often a patient refuses further treatment ostensibly because it is toxic, but in truth perhaps because it is inconvenient or because he decides he feels better and therefore needs no further treatment. Thus we have many causes of patient loss. Some of it is merely apparent loss because we have received inadequate data but at least have a chance of getting the data. Sometimes patient loss is due to death from other causes—accident, suicide, other diseases—and we are left with the problem of deciding whether suicide may mask drug failure or a toxicity effect. Are the "other diseases" truly "other diseases" or are they some type of toxicity?

Finally, the questions raised by Dr. Zelen are fascinating. Recently we submitted to an Ethics Committee a protocol that involved current best treatment in AML; we were comparing Ara-C and daunorubicin with Ara-C and a different dose of daunorubicin and Ara-C and Adriamycin. We were brought up short by the Ethics Committee who said to us, "If this is the current best treatment, how can you give something else?" Taken aback, we replied that that was the basis of the study. We knew some patients have responded to these other regimens or to these other doses; we were trying to establish whether the "current best" treatment was really the current best treatment. We finally got them to approve the study, but I'm not sure we could get them to approve Dr. Zelen's approach. The people on the committee who gave us trouble were an endocrinologist, a lawyer, and a person who was neither a physician nor a lawyer. Once a single person raised the question, the whole committee developed doubts.

**Dr. DeWys:** Dr. Zelen, in the last part of your presentation, when we got to the design in which the patient made the choice, if the patients choose A and B with equal frequency, you have a $P$ value of 0.5 (or 0.25 when you square that), so that you have an efficiency of one fourth of the usual design; that is, you need four times as many patients. That seems almost self-defeating. In that kind of setting, would it be possible to take the patients in the G2 group who accepted treatment A, compare them with the treatment B group, and in that way reclaim some efficiency?

**Dr. Zelen:** No, because of the possibility of built-in

bias. One has to weigh the "statistical efficacy," which in this case, as you pointed out, is quite low, with the ethics of doing the trial and with the option it gives patients to select what is appropriate for them. One may have to have four times as many patients in the trial in order to run a study when the physician does not believe that both therapies have equal therapeutic benefit. Furthermore, such studies may bring in many more patients, and in the end may represent a more efficient way of conducting studies.

Every patient has a right to the best standard treatment. When a patient enters an institution, the physician does not discuss with him all the therapies that are under investigation in that institution—only those therapies which the physician believes appropriate. In trying to extend this privilege of treatment choice, we can do so on a lottery basis, with no reference to the patient's personal characteristics or economic status, and on a 50–50 basis. In current practice, many physicians who participate in randomized studies may not tell the patients exactly what randomization means; that is, informed patient consent does not necessarily mean that one describes to the patient what is happening. As one surgeon said to me recently, "Why, if I told my patients how their therapy was really selected, I'd have no patients in a randomized study." If the Department of Health, Education, and Welfare (HEW) modifies its guidelines to make them more restrictive, randomized studies may come to an end in this country. In the meantime, we must make recommendations concerning how randomized studies can be carried out, and we must be sure that when patients are offered experimental therapies, they are fully informed of this fact.

**Speaker from the floor:** Dr. Zelen seems to have described what might be called randomized concurrent historical control. As statisticians and clinicians we are challenged to find a way of combining brilliant statistical methodology and clinical reality. Dr. Zelen's plan brings up all sorts of ethical issues and stimulates us to consider certain fundamental ethical issues. Unfortunately, there is little general agreement on the criteria of an ethical issue except perhaps that a study is ethical if the Human Research Committee approves it. Regarding informed consent, my second-year medical students are enthralled with the concept, which they believe brings all sorts of good things to the patient. Then I ask, "How do you obtain an informed consent?" "What kind of physician will you be if you can't get a patient to agree to anything that you tell him is reasonable?" Perhaps we may be doing more harm than good with informed consent.

**Dr. Simon:** An earlier comment was made concerning the patient characteristics in the proportional hazard model. We do not need procedures for testing the validity of all models, and I believe the models put forward here provide a very good fit to the data. However, it should be recognized that this is not completely a science; you might find, for example, that different sets of three variables would perhaps provide an equally good fit to the data.

Concerning the question of ethics, if everybody is offered new treatment, and if you think the new treatment might benefit the patients, everyone is offered that option in the non-randomized study. I am not sure I agree with Dr. Zelen that there are limitations in the numbers of patients. Sometimes this may be true, but in many situations investigators are looking for patients and it is hard to get patients to participate.

**Dr. Zelen:** The type of surgery was determined chiefly by the stage of disease. Other sets of three factors that could fit these data—a set including type of surgery—could have been included, but if it is highly correlated with stage, it would not have made any difference in the outcome. It could have been included but was not because it was determined by stage of disease.

**Speaker from the floor:** Dr. Zelen, you say that the stage of the disease is determined after surgery. You also have the clinical stage, the type of operation, the pathologic data. You did not comment on any of them. Presumably the type of surgery is determined by stage. One would assume it would have to be based on clincial assessment and prognostic variables from pathologic stage.

**Dr. Zelen:** All our staging is according to the UICC classification; that is, it is clinical pathologic staging. But if the patient has extensive skin involvement from the beginning, you know from direct observation that he has stage 3 disease, but large nodes in the axillae allow you to make a clinical diagnosis of stage 3 disease.

**Speaker from the floor:** In that case, why use nodal status rather than clinical status as a variable?

**Dr. Zelen:** This particular decision was based on a combination of the clinical size of the tumor, skin involvement, and whether the nodes were positive or negative. Patients with negative nodes would not have been included even though such patients might have stage 3 disease, because we did not believe that we were justified in including them.

**Speaker from the floor:** My purpose was to point out that your paper did not discuss a detailed analysis of the application of surgery as the primary treatment variable.

**Dr. Zelen:** From a clinical point of view, surgery or any local treatment, even radiation therapy, if it is effective, achieves very comparable local regional control when given following "lumpectomy." That is why we did not emphasize that point because surgery is done only for local regional control, and we were talking about disease-free survival and overall survival.

**Dr. Moffatt:** Dr. Zelen, what would happen if most of the patients did not choose treatment B, the experimental treatment? One of the reasons for randomizing the patient's treatment is that it avoids bias in the selection of patients or in the type of treatment. Would you not open

the study to such bias by allowing the physician to inform the patient, who is then encouraged to choose the therapy *he* wants?

**Dr. Zelen:** It is much easier to say "Would you like this experimental treatment?" and then describe it rather than describe both treatments and discuss the relative merits of each. If you mount a study and most patients elect not to receive treatment B, I would conclude that it was premature and that the patients may be acting in their own best interests. With regard to randomization, this design has the merit that treatment is known in advance of consent. Because initially you divided the patient population into two groups, you have two comparable groups, you compare one group to the other, and there is no bias. It does lose something if there are any interactions which exert a positive or even negative effect on the therapy. However, most of the studies we deal with contain these inherent biases as to who comes into the study, and many of the trials do not have sufficiently large numbers of patients to permit us to explore interaction effects with any sensitivity.

**Dr. Moffatt:** Since you selected certain patients to study initially and randomized to determine who would be given the privilege of choosing treatment, I do not see what you gained by allowing such randomization at that stage.

**Dr. Zelen:** You are comparing two groups: one group that we have treated in the same way; another group that may be treated by two therapies, depending on the selection process. If the two groups are comparable with respect to the endpoints you are investigating, you might conclude that use of the experimental treatment in a portion of the patients did not affect the collective outcome of group 2. On the other hand, if the outcome was different, you would conclude that the only difference between those groups was the use of treatment B for some individuals.

**Dr. Andrews:** How does optimization relate to what has been presented here today? Is it considered to be something quite apart? Can it be studied within the framework of the clinical trial, and if so, how is its study approached? I work in radiation therapy and believe that we are close to the point where we can optimize the use of that modality. We speak loosely about the "best" treatment, but seldom if ever offer evidence that it may be "modality effective."

**Dr. Zelen:** No, we are not talking about optimization today. We are talking about one part of the science of therapeutics. We are addressing the following question: When methods get to the point of testing in humans, or when they have received some tests and you want a definitive comparison of some new method to the standard methods, what study will give definitive and reliable answers?

**Dr. Margolis:** In the context of the example Dr. Zelen used, namely, mastectomy, it is technically true that women are being deceived if they are not told that they could have received another type of therapy. What would these patients say if they found out? There are 60,000 or 70,000 breast cancer cases a year, of which probably 50,000 or 60,000 are primary stage 1 and stage 2. As chairman of the Segmental Mastectomy Protocol in the NSABP, it bothers me that of 50,000 or 60,000 eligible cases only 2000 have been entered into the study in 2 years. What are the other 58,000 patients going to say when they discover that their surgeons directed them into a mastectomy or a form of total mastectomy without informing them that the other option was open? A very important question is whether you can cure cancer of the breast without removing the breast. I think medical ethics demands that we get this answer. The current situation with respect to informed consent is preventing us from getting this answer for our patients. We are not serving our patients as well as we should as physicians and scientists. It does not matter what kind of statistical assessment you do of the data we have now; we need Dr. Zelen's kind of imagination. His format does need more thought, but we have to break away from the rigid confines of current approaches if we are going to answer this question.

**Dr. McGill:** Dr. Gehan, I believe I see an almost architectural flaw in your study. You have a historical control group which you compared to an FAC-BCG treatment, and you concluded that the way to treat these patients is with FAC-BCG. However, it is not clear that your historical control group received an equal amount of therapy, either with Adriamycin, Cytoxan, or some other combination.

**Dr. Gehan:** The control group received the surgery and radiation; it did not receive any adjuvant treatment. The important point is not that FAC-BCG is the best way of administering adjuvant chemotherapy, but that FAC-BCG is good adjuvant chemotherapy. In fact, some of our results with FAC-BCG are not very different from those reported by Dr. Bonadonna. The pressing question in breast cancer right now is what the best form of adjuvant chemotherapy is. To summarize, we are saying that FAC-BCG is a good treatment. Furthermore, a study done at M. D. Anderson Hospital shows that FAC-BCG delays the onset of the second recurrence as well as that of the first; however, many patients are not cured by FAC-BCG.

**Dr. McGill:** Dr. Thompson or Dr. George, I think you are arguing in favor of sequential trials in clinical oncology, yet the data you presented were essentially from an infectious disease trial with a close endpoint. Is that the reason why sequential trials are not used—they are less applicable in oncology, where we must examine delayed endpoints and multiple endpoints?

**Dr. George:** I think you are right. One of my points was that the methodology of followup trials is not as well-developed as the more straightforward sequential clinical

trials. Whether we say so or not, many trials that are being analyzed in cancer are indeed being done in a way that would qualify them for the label "sequential"; that is, they are being done several times before the results are published. For the most part this fact is ignored. I believe it is wrong to ignore it, and it is obvious that the ordinary *P* value significance level is no longer relevant in its original statistical sense.

**Dr. Fihos:** I suppose about 200–300 clinical trials are going on now. Has anyone done a study to determine how good these clinical trials are? In other words, are the statisticians doing their jobs?

**Dr. George:** I do not believe that the discovery of a cure for cancer has been delayed because statisticians are not doing their jobs.

**Dr. Fihos:** A study published in the British Journal of Cancer suggests that > 50% of all clinical trials have some defect in their design. The statistician would perform a great service to the physician if he could devise some guidelines to indicate that a given technique is wrong for a given purpose.

**Dr. George:** Every clinical trial can be improved. However, in all of these large institutional studies there are many compromises between most of the major figures involved. There is also a compromise with review committees, with permission to mount trials, and often the fundamental science is not well understood by the people proposing the trials. Therefore, one should not judge these trials by the statistics; one has to judge them perhaps by their methodology, by how successful they have been in attaining their objectives. Frankly, we are just learning how to judge trials because running clinical trials is the most difficult way of generating scientific data. Whenever we can, we should generate the necessary scientific data in the laboratory. However, the science of therapeutics is developing empirically, and this fact commits us to the clinical trial method.

**Member of UICC Committee on Controlled Therapeutic Trials:** I would like to comment on that. The UICC has a project concerned with assessment of clinical trials in terms of the impact on actual treatment. The study is not yet complete, but Dr. Kenis from Brussels reviewed the literature on lung and breast cancer, and I believe Dr. Staquet reviewed the literature on colorectal cancer. Their paper suggests that the major impact has been to encourage more quantitative collection of data and closer adherence to the protocol. This has led to important changes. For example, in lung cancer, surgeons now tend to do lobectomies rather than pneumonectomies, yet there has never been a randomized trial of lobectomy. In breast cancer, there is now a greater tendency to do simple rather than radical mastectomy, yet there has not been a randomized trial of simple versus radical. It is the kind of data collection acquired through clinical trials that has led to these changes. In colorectal cancer, postoperative radiation has been shown to be beneficial in the randomized study, but many of the radiotherapists don't administer it that way. In the long run the real test is the confirmatory studies. Reviews of the literature directed at an assessment of the impact on treatment are difficult but laudatory.

**Speaker from the floor:** Dr. Jacobs, the NSABP did a study which controlled the simple mastectomy against the radical. It was a pure study with clinical stage 1 tumor: total mastectomy, total mastectomy and radiation, and radical mastectomy.

**Dr. Jacobs:** Dr. Zelen, to an increasing degree, the news media are giving tremendous exposure to cancer and new cancer treatments, and when patients are in a clinic or a private doctor's office they talk to each other about new treatments. Thus a major problem in the randomized group 1 is that many of these patients have heard about treatment B before treatment B is even available to the physician. In that event, what happens to those patients who get randomized to group 1 and then ask about treatment B?

**Dr. Zelen:** That is a serious problem. Clearly if a patient comes in and asks about a particular kind of therapy, the physician is obliged to tell him or her about that therapy. If too many patients in group 1, say, learn about experimental therapy group B and elect it, we have a null study. Recently a science reporter from the *Washington Post* told me he had been doing a story on a trial of simple versus radical mastectomy for breast cancer, and in the process, had been investigating patients' consents. He interviewed three patients. One patient had declined to enter; the other two patients had no recollection of ever having been told that they were randomized, yet all documents show that they had been informed. Bill Curran, a professor of legal medicine and an authority in this area, believes that there is no legal problem in implementing the approach I have described. Other legal opinions have been almost unanimous that my plan involves no compromise or violation of existing HEW guidelines.

**Dr. Volger:** One of our problems in conducting clinical trials, particularly randomized studies, has been the misinterpretation of interim analyses. We are often misled by the fact that we gather a *P* value that seems significant and then, as time goes on, it seems to drop. Dr. George or Dr. Sedransk, do you believe that the procedures you are pursuing will be of some value in reducing such misinterpretation?

**Dr. George:** The issue of interim analyses has received considerable attention as a theoretical construct, but we have not looked at what happens in practice. I am speaking of cooperative trials, of course. Many people have learned to distrust the interim analyses because of missing data, and ignore them in terms of realistic decision-making. Thus what may have been a trial that embraced the possibility of an early stop if it was warranted becomes a trial that will never be stopped early, because

no one believes the interim analyses. The trend today is not to be misled by the early results but simply to ignore them. The answer is, yes, if you have all the data, sequential procedures do help.

**Dr. Sedransk:** The procedure I am pursuing operates on the principle that it doesn't depend less on the order in which the data are acquired or how many times you look at them than on your own sets of patients. The notion that the *P* values drop results from survival-curve analyses where the extrapolations that are used tend at the outset to be unduly optimistic. This fact is demonstrable, and it is due to the censoring problem. There is no way to avoid the censoring problem entirely; if the information is there, none of these methods is going to bring it in faster, but they will make better use of it. The methods that Dr. George and I have talked about can be dealt with quite directly. Both methods are designed to avoid trouble with multiple testing and with the discrepancy between the *P* values one quotes and the *P* values that actually function.

**Dr. George:** May I discuss one other problem? Early on, you have two remission duration curves and they look very impressive, but later in the trial they no longer look so impressive. One explanation is that if they do not look impressive early, nobody reports them. The ones that are very impressive early are the ones your eye is drawn to and the ones that people present, and, by the laws of chance, many of these differences are spurious.

**Dr Cantril:** I represent practicing community physicians. Patients come to us as physicians and look to us for advice. To a greater or lesser degree some patients want to participate in the decision-making process and some do not; they have already decided to come to the physician, and they assume that he will make all decisions for them. I believe we have a serious problem when an informed physician sees two or three alternative approaches to cancer therapy and does not inform the patient of his options.

**Dr. D'Angio:** Perhaps the answer is obvious and I am just not seeing it, but I am troubled because I don't understand. I believe medical ethics requires that a patient be informed *before* he is entered into a study, that is, before he is randomized. The consent, then, is consent to be randomized so that the information is "upstream" before your patient is then randomized to G1 or G2 in your scheme. Are you not obliged to tell the patient that he is entering into a trial and, therefore, are not saving anything at all?

**Dr. Zelen:** No, you are obliged to tell the patient only when you deviate from normal practice. If the patient is getting the best standard therapy (I am now reflecting legal views), that is his expectation and his right. When you deviate from that, the patient has to be informed. With regard to using data from the patients in the so-called control or best therapy group, such data are used all the time in any clinical department to assess therapy without the patient's knowledge. This happens in tumor registries, with periodic reviews, retrospective studies, and case-control studies. The important thing is to protect the patient's privacy so that the Privacy Commission agrees that you are using the data for statistical purposes.

**Dr. D'Angio:** I agree that as long as this patient is going to a physician and getting the best treatment possible, say for stage 1 carcinoma of the breast, and the physician believes a radical mastectomy is the best thing to do, there is no problem. However, if the patient walks in the door and the physician recognizes that she is now going to be entered in some kind of study, even though one of the randomizations is to "best standard" therapy, she is now engaging in a study and ought to be told that.

**Dr. Zelen:** It is hard to say what is correct and what is not correct, because ethics change by the moment. It depends on the physician's personal experience. I believe that what is happening now may not reflect the HEW guidelines for informed patient consent; indeed, this term does not necessarily mean that the patient understands what is going on. I am proposing a way to make the decision processes simpler, so as not to anger the patient by informing him that his treatment will be chosen by tossing a coin, to have situations where the two therapies may not have identical outcomes but may involve other factors, such as the quality of the life or the kind of damage associated with treatment. As we well know, many men would refuse to have genitourinary surgery because they do not want to become impotent, so that individuals consider all kinds of utility functions. I would like to incorporate these factors into the machanisms by which people are entered into experimental therapies.

**Dr. Kennedy:** How reliable is the final product? If a patient walks into the doctor's office and the doctor goes through the process of selection, some patients will be entered into a study eventually. However, most of those studies do not log in all the patients who came into that investigator's office. How much bias does this subtle selection produce in your final end product? A good example is the advanced breast cancer patient entering into five-drug chemotherapy. Obviously many of those patients have advanced disease, while an early recurrent cancer might have been excluded and viewed in a different way. Eventually, the investigative institution will publish results which may have an important impact on practice, but a large group of patients have been excluded at the outset. The old Breast Cooperative Group did log in every patient the investigators saw, but I do not think most studies do that now. Do you think this process produces a bias and perhaps misdirects some of our final conclusions?

**Dr. Sedransk:** The degree of bias will depend on how different the group of patients treated in the study is from the group to which the results are being interpolated. The best protection against being misled by such

bias is to describe very clearly the group of patients treated in the study so that it is clear to whom these results can be interpolated. The notion of logging in all patients contains the inherent difficulty that over a period of years, patient referral patterns change even within a given institution and certainly differ among institutions. All breast cancer patients coming to one institution are not necessarily the same population as all breast cancer patients coming to another institution. All that logging in would accomplish would be to identify all such patients for a single institution. In and of itself, logging in is not particularly useful; at least, it is not as useful as clearly describing those patients who were studied and trusting the reader to make only legitimate extrapolations of the study results.

**Dr. Zelen:** I would like to comment on that. I believe logging in the patients would raise the quality of the interpretation of the final results. We did a study some time ago on 18 ECOG studies in which the same therapy was used in back-to-back trials. In many of these—same group, same institutions, same therapy—the differences were statistically significant. Depending on how things work out, institutions adjust their methods for the number of patients who come in. Unfortunately, the National Cancer Institute has not regarded logging in as very important and, as a result, no funding has been provided to do it. It would be a big job, but if we had the resources, it would raise the reference frame concerning to what our results can be extrapolated. I am in favor of it.

**Chairman:** Finally, Dr. Ellison asked how many stratifications there should be. There should be at most eight.

# Operative Procedures: Is Standardization Feasible or Necessary?

**Jerome J. DeCosse,[1,*] William L. Donegan,[1] Neil Sedransk,[2,3] Neil F. Jones,[1] and Dann B. Claudon[4]**

## SUMMARY

**Standardization of surgical procedures is both feasible and necessary in clinical trials. To the extent that an operation may influence prognosis, identify prognostic factors, or reveal variables upon which additional treatment is contingent, the operation must be defined precisely enough to permit the determinations to be made uniformly and to enable the surgeons to meet the necessary standardization. If uniform procedures are to be applied widely, surgeons who participate in therapeutic trials must meet the standards of clinical scientists and accept special obligations for education and self-discipline.**

[Cancer Treat Rep 64:419–423, 1980]

Reviews of combined treatment modality protocols generally show meticulous attention to the details of drug therapy and careful quantitation of the physical determinants of radiation therapy, but often only a cursory description of the operative procedure. The title of the procedures is deemed sufficient direction to permit all surgeons to perform a uniform cancer operation.

It is common knowledge, however, that surgeons purporting to do the same operation perform it in a variety of ways. Operations of similar title embrace surprisingly wide variations of technique, so that a "standard" operation may require 1 or several hours to complete and may produce a variety of surgical specimens. It is difficult, however, to obtain accurate information on the results of such surgical license in terms of tissues removed and any potential influence upon prognosis.

To assess the comparability of tissues removed by different surgeons performing a "standard" procedure, we counted the number of lymph nodes in surgical specimens obtained during a consecutive series of operations described as radical mastectomy or modified radical mastectomy. Radical mastectomy for cancer of the breast was chosen as a model because the operation is probably defined more precisely than any other cancer operation. Furthermore, this operation is performed so frequently that it is part of the formal training of all general surgeons.

## MATERIALS AND METHODS

During the 4½ years ending in June 1977, 143 patients underwent a radical mastectomy for cancer of the breast at Columbia Hospital (Milwaukee, WI). These operations were done by 15 certified general surgeons, six of whom performed at least ten operations each. During the same period, 128 patients were submitted to procedures described as a "modified radical mastectomy." Throughout the study the method of processing specimens in the surgical pathology laboratory remained uniform. The fresh specimen was sectioned serially and all lymph nodes encountered were removed, enumerated, and examined microscopically. Individual surgeons were compared to each other anonymously with respect to the number of lymph nodes in the specimens they submitted after radical and modified radical mastectomies.

The Kolmogorov-Smirnoff method was used to test differences; a $P \leq 0.05$ established significance.

## RESULTS

### Trend in Operations

During these 4½ years the proportion of modified radical mastectomies to radical mastectomies increased

[1]Department of Surgery, Medical College of Wisconsin, Milwaukee.

[2]Division of Statistical Science, State University of New York at Buffalo.

[3]*Present address:* Department of Mathematics, State University of New York, Syracuse.

[4]Department of Pathology, Columbia Hospital, Milwaukee, WI.

**Reprint requests to:* Dr. Jerome J. DeCosse, Department of Surgery, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10021.

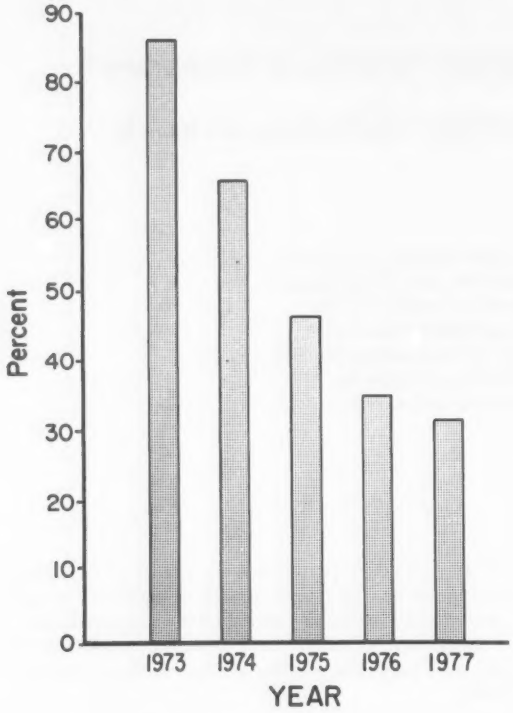



FIGURE 1.—Radical mastectomies as the proportion of all radical and modified radical mastectomies for 1973–1977.

steadily. Modified radical mastectomy accounted for 16.4% of all radical and modified radical mastectomies performed in 1973 and 65.8% of those performed in 1976 (the last full year of the study) (fig 1).

## Age and Presence of Metastases

The number of lymph nodes found in specimens from both operations did not appear to be affected by age except in patients who were > 80 years old. A slight decrease was noted in specimens from women in their 70s, and this decrease was more pronounced after a modified radical mastectomy. The presence of metastatic cancer in axillary lymph nodes did not influence the total number of lymph nodes encountered in either radical or modified radical specimens. In the latter specimens, however, the range was wider when all lymph nodes were negative.

## Year of Surgery and Number of Lymph Nodes

For radical mastectomy, the median number of lymph nodes found in the specimens increased slightly from 1973 to 1975 and then declined (fig 2). For modified radical mastectomy, the median number of lymph nodes declined slightly during the study period, whereas the range increased (fig 3).

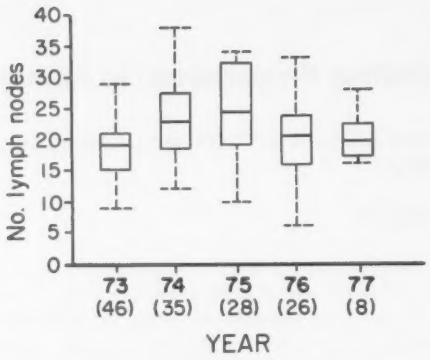



FIGURE 2.—No. of lymph nodes found in radical mastectomy specimens by yr: horizontal bar = median, box = middle 2 quartiles; and dotted lines = range. Nos. in parentheses = No. of radical mastectomies.

## The Surgeon and the Operation

The number of operations performed by the surgeon had no effect on the median number of lymph nodes found in radical mastectomy specimens (fig 4). No difference was noted among surgeons with respect to the number of lymph nodes removed. However, surgeons varied in the range of lymph nodes encountered in their radical mastectomy specimens.

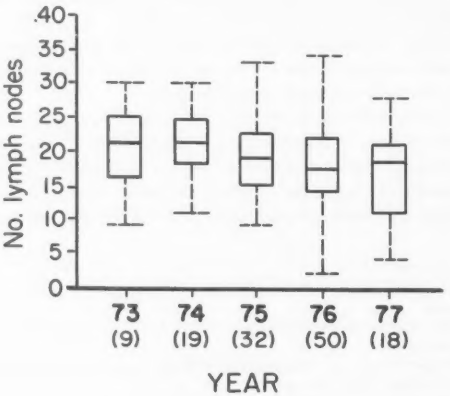



FIGURE 3.—No. of lymph nodes in modified radical mastectomy specimens by yr. Nos. in parentheses = No. of modified radical mastectomies.

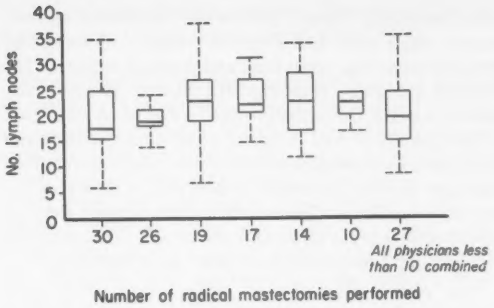



FIGURE 4.—No. of lymph nodes in radical mastectomy specimens by surgeon. Each vertical bar except extreme right = 1 surgeon.

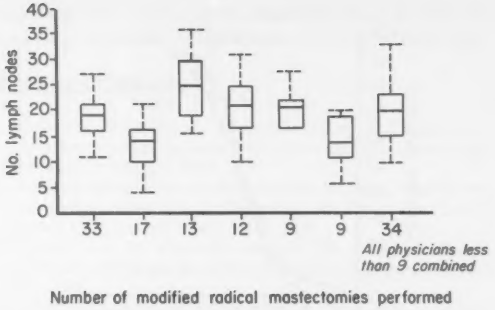



FIGURE 5.—No. of lymph nodes in modified radical mastectomy specimens by surgeon. Each vertical bar except extreme right = 1 surgeon.

The median node counts were between 18 and 23 for surgeons who performed at least ten radical mastectomies. All other surgeons together had a median number of 20 nodes; the range of lymph nodes in specimens submitted by this latter group (nine surgeons, totaling 27 operations) was no wider than for two surgeons who did 19 and 30 mastectomies, respectively. Hence, the specimens submitted by surgeons who performed large numbers of mastectomies exhibited as much variability as the specimens from those surgeons who performed few mastectomies.

In modified radical mastectomy, the number of operations performed by a surgeon did not determine the number of lymph nodes in his specimen (fig 5). Even among physicians performing at least nine operations, the median number of lymph nodes varied from 14 to 25.

Surgeons who performed radical mastectomy predominantly had more lymph nodes in their radical specimens than in their modified radical specimens (fig 6). In contrast, the specimens submitted by those surgeons who did chiefly modified operations contained the same number of lymph nodes regardless of which operation had been performed (fig 7).

The radical mastectomy specimens from surgeons who usually did the radical procedure contained more lymph nodes than the specimens of those surgeons who usually did modified operations. Yet, modified operations of the former contained fewer lymph nodes than did those of surgeons who preferred modified mastectomy.

## DISCUSSION

This study identified a trend toward the modified radical operation during the 4½-year period. The review indicated that radical mastectomy was a relatively consistent operation when judged by the median number of axillary lymph nodes obtained from surgical specimens, although there was considerable variation in the range of nodes obtained by individual surgeons. The variation observed could not be ascribed to the experience of certified surgeons, age of the patients, or presence of nodal metastasis.

As might be expected, there was a wider variation in the number of axillary lymph nodes removed during "modified radical mastectomy"; the several types of modified procedures practiced differ chiefly in the extent of axillary dissection. It was noted that the yield of lymph nodes in similarly designated operations differed, depending on whether the surgeon's usual preference was radical mastectomy or modified radical mastectomy.

Although no evidence is offered that these differences affect prognosis, these findings support the utility of standardization in clinical trials, and perhaps could be generalized to operations for cancer at other sites.

For the purpose of clinical trials, the standardization of cancer operations must be considered a desirable goal, and, with effort, an achievable one. The purpose of clinical trials is to answer well-defined questions regarding the clinical management of cancer. In its simplest form the objective is to detect a significant change in a de-

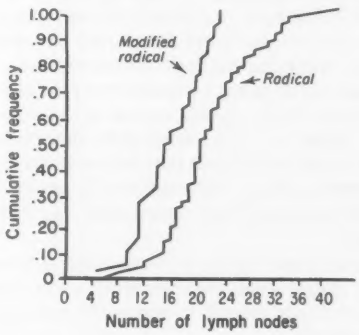



FIGURE 6.—Cumulative frequency plot of lymph nodes by surgeons doing predominantly radical mastectomy.

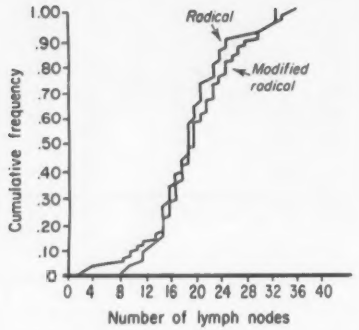



FIGURE 7.—Cumulative frequency plot of lymph nodes by surgeons doing predominantly modified radical mastectomy.

pendent variable, eg, 5- or 10-year survival, disease-free survival, or some other prognostic determinant, following a change in one independent treatment variable. If the operation is not the independent variable and the quality of the operation could influence the outcome of treatment, the procedure must be maintained as nearly constant as possible. Laxity in this regard constitutes a breech of the scientific method.

There is little doubt that standardization of operations can be improved through education and surveillance. A detailed description of the surgical procedure, including options and limitations, is an appropriate part of a research protocol. Protocols should resolve uncertainty by identifying operative alternatives and indicating either inclusion or exclusion from study as may be appropriate to study design. In certain cooperative groups, internal Surgical Quality Control Committees have been set up to educate surgeons in advance of participation, in part by developing and disseminating illustrative and descriptive material regarding the surgical aspects of the protocol. Reinforcement is provided through continuing education. These committees also review and make policy decisions regarding protocol deviations, and, if necessary, visit participating institutions to resolve problems. Surveillance should include review of the surgeon's operative report and the pathologist's description of resected tissues to ascertain that both conform to protocol requirements. To ensure compliance, it may be necessary to require a photographic record of an operation and the specimen.

Standardization of operative procedures is necessary first, to ensure a uniform therapeutic effect (however it is measured) and second, to ensure the accuracy of staging procedures.

It is surprisingly difficult to demonstrate whether minor or even major differences in the extent of an operation affect prognosis. Since it is believed that cancer spreads from a single focus and since the range of its extension remains uncertain, wide local and regional resection has been an attractive principle. The paucity of convincing data to support this concept may be due in no small part to the dearth of well-designed randomized studies and the weakness of retrospective reviews in which undefined selection factors determine a nonrandom choice of operative procedures. Clinical trials which suggest that the extent of surgery is important are usually confounded by concomitant nonuniform treatment interventions (1).

Less controversial than its impact upon prognosis is the clear utility of surgery to pathologic staging. Staging based upon resected tissues has long had an important role in determining prognosis, and, in recent years, has had a major influence in treatment planning for Hodgkin's disease and in identifying patients for postoperative adjuvant chemotherapy. Pathologic staging ensures the comparability of clinical material, but its credibility depends upon consistency, timing, and a close link between the surgeon and the surgical pathologist.

The scope of an operation can provide important information in one case which is not available in another because a different or a less complete procedure was performed. For example, if the presence of metastases in axillary lymph nodes determines the selection of further treatment or is a criterion for randomization, the axilla must be sufficiently sampled to establish the presence or absence of metastases. Without further definition, however, the nature of sufficient sampling remains controversial (2). If an abdominal operation is included as therapy in one arm and reveals the presence of occult pathology, eg, hepatic metastases (a discovery that is not possible in the other arms), such cases may be excluded from one treatment arm and not from another.

If the primary function of an exploratory operation or a resection is staging, other invasive diagnostic procedures, such as laparoscopy, liver biopsy, or lymphangiography may also contribute information and influence the comparability of clinical material. Protocols should be designed to accommodate these uncertainties by specifying the range of diagnostic procedures permitted and by defining their influence upon the design of the study.

The sequence of an operation in multimodality treatment protocols can also affect the comparability of treatment groups and the validity of results. Lingering doubt that comparability was achieved has troubled randomized trials in which radiation therapy or chemotherapy before operative intervention was associated with a preponderance of favorable neoplasms compared to the group that had surgery alone (3). This apparent "down staging" may reflect the effects of treatment or an inequitable distribution of cases; one's decision on this point has an important bearing upon interpretation of the results. Uniformity in the sequencing of surgery for therapy or for

staging provides further assurance of comparability and deserves attention in the design of clinical trials.

The validity of surgical staging depends upon a uniform examination of resected specimens. Subsequent randomization and stratification of patients are often contingent upon the pathologist's evaluation. An operation of proper extent can easily appear deficient in the light of a casual report from the pathologist. Furthermore, a diligent search of resected tissues may produce 25 lymph nodes for examination and a casual search may produce only five. Although there are reports to the contrary (2,4), one must assume that the pathologist can learn more from examining 25 lymph nodes than from examining five; indeed, much less evidence is obtained from examining a single section of a lymph node than from examining multiple sections (5,6). Thus, the consistency and quality of the surgical contribution to a clinical protocol are dual responsibilities, and require standardization in both the operating room and the department of surgical pathology.

## REFERENCES


1. Atkins H, Hayward JL, Klugman DJ, et al. Treatment of early breast cancer: a report after ten years of a clinical trial. Br Med J 2:423:429, 1972.
2. Roberts MM, Blumgart LH, Davies M, et al. Simple versus radical mastectomy. Preliminary report of the Cardiff breast trial. Lancet 1:1073–1076, 1973.
3. Higgins GA, Jr, Conn JH, Jordan PH, Jr, et al. Preoperative radiotherapy for colorectal cancer. Ann Surg 181:624–631, 1975.
4. Fisher B, and Slack NH. Number of lymph nodes examined and the prognosis of breast carcinoma. Surg Gynecol Obstet 131:79–88, 1970.
5. Pickren J. Significance of occult metastases. Cancer 14:1266–1271, 1961.
6. Saphir O, and Amromin GD. Obscure axillary lymph node metastasis in carcinoma of the breast. Cancer 1:238–241, 1948.

# Quality Assurance Programs in Clinical Trials[1]

Arvin S. Glicksman,[2,3,*] Lawrence E. Reinstein,[2,3] Roger Brotman,[2] and Daniel McShan[2,3]

## SUMMARY

**Quality assurance programs attempt to assess the uniformity of compliance with a stated study program so that the results from one institution will indeed be comparable to those of all of the other institutions in the group. In the best situation, it should be easy to compare results from one group to those of any other group and any other studies of the same disease, because the basis of selection of the patient population, the treatment of patients, and the endpoints have been uniformly defined throughout. The evaluation of the end results includes a quantitative assessment of the appropriateness of these parameters. Quality assurance programs document the validity of interinstitutional, intergroup, and international studies. They are time-consuming and costly. They can be ulcerogenic. But, in the end, they are the firm basis upon which the statistical analysis can proceed.**

[Cancer Treat Rep 64:425–433, 1980]

The validity of clinical trials performed under a single protocol in multiple institutions is predicated upon the premise that the selection of patients and their treatment will be uniform at all the participating institutions. This assumption requires a concise definition of the population to be studied, the treatment regimens to be followed, and the methods of evaluating the results. Quality assurance programs are an attempt to document the validity of the assumptions and to quantitate the extent of variance.

## PATHOLOGY REVIEW

The first question to be answered is whether the diagnosis is appropriate for the protocol. Without pathology review panels, this cannot always be certain. One of the first successful attempts at pathology review was made by the Lymphoma Panel. The Lymphoma Panel has not had a serious discordance between the local pathologists and the central panel in establishing the diagnosis of Hodgkin's disease, although there has been appreciable shuffling about the type of Hodgkin's disease. This was not a major problem in the Canadian-American Study on Hodgkin's Disease because there were no differences in treatment options based on pathology. In the new study by the Cancer and Leukemia Group B (CALGB) on Hodgkin's disease, CALGB-7751 (the comparative effectiveness of combination chemotherapy alone and with radiation therapy by involved field or extended field, in poor risk patients with stage I and II and Hodgkin's disease) which is now accruing cases, this has become a matter of concern since patients with lymphocyte-predominant and nodular-sclerosing Hodgkin's disease without symptoms, ie, *A* classification, are excluded from these studies. The same situation pertains to the new study by the Southeastern Study Group (SEG), SEG-77HD0103 (phase III study of combination chemotherapy with BCNU, vincristine [VCR], cyclophosphamide [CTX], prednisone [PRED], and procarbazine with or without involved field radiotherapy in early-stage Hodgkin's disease with unfavorable prognostic indicators), which has also been implemented.

The Lymphoma Panel and local pathologists disagree as often as not in establishing the subclassification of non-Hodgkin's lymphoma. The Southwest Oncology Group protocol, SWOG-7425 (chemotherapy with CTX/Adria-

[1]Supported in part by Public Health Service grant CA-21058 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare (through the Cancer and Leukemia Group B Radiotherapy Committee Office).

[2]Department of Radiation Oncology, Rhode Island Hospital, Providence.

[3]Section on Radiation Medicine, Brown University, Providence, RI (for the Radiotherapy Committee of the Cancer and Leukemia Group B).

**Reprint requests to:* Arvin S. Glicksman, MD, Department of Radiation Oncology, Rhode Island Hospital, Providence, RI 02902.

mycin [ADR]/VCR/PRED/bleomycin versus CTX/ADR/VCR/PRED and immunotherapy with bacillus Calmette-Guérin versus CTX/VCR/PRED/bleomycin for non-Hodgkin's lymphoma) and the SEG-76NHL0103 (phase III study of radiotherapy followed by combination chemotherapy with CTX/VCR/PRED for non-Hodgkin's lymphoma, stages II and III) did not have separate therapeutic arms based upon histology and therefore were able to utilize all the material by retrospective stratification. In CALGB-7651 (chemotherapy with VCR/STREPTONIGRIN/PRED and VCR/PRED/CTX plus radiotherapy for stage III and IV lymphocytic lymphoma [lymphosarcoma])and CALGB-7652 (radiotherapy and chemotherapy with VCR/PRED/STREPTONIGRIN, VCR/PRED/CTX, and VCR/PRED/ADR for histiocytic lymphoma [reticulum cell sarcoma]), treatment options were based upon histology. Patients diagnosed as having lymphocytic lymphoma at their local institution were randomized between radiotherapy and no radiotherapy intensification after induction chemotherapy. The patients with histiocytic lymphoma, on the other hand, had a three-way randomization between no intensification, ADR intensification, or radiotherapy. Because the Lymphoma Panel's diagnosis became known 2–3 months after patient registration, one third of the patients had to be excluded because of the change from histiocytic to lymphocytic or vice versa.

Pathology review of solid tumors has not been as well-detailed, and yet many protocols are being done for lung cancer based upon histologic type. There are a number of protocols for adjuvant chemotherapy in breast cancer based upon the number of positive axillary nodes, yet no standardized review of the pathology techniques for the clearing of those nodes in the axilla has been built into the studies in all groups.

It is clearly essential, therefore, that where treatment options are predicated on specific subsets of pathology, a central review of the histology be performed in real time, so that the assignment to treatment arms based on histology can proceed without delay or exclusion of an appreciable number of entered cases. It is equally essential that criteria for the specific diagnosis be established and remain consistent throughout the performance of the study. Changing attitudes by the pathologist mid-trial can be disastrous. This can be overcome either by again reviewing all the cases already entered or by deciding to maintain the original criteria throughout. What is more difficult is to compare studies of different years if changes in criteria have become established in the interim.

An increasing number of breast cancer protocols are stratifying patients on the basis of the presence or absence of hormone-receptor sites in the tumor. Yet the technique for this analysis is far from uniform. The amount of fat or connective tissue included in the specimen and the rapidity of preparing the specimen in liquid nitrogen, as well as the lab techniques themselves, can all affect the results of this analysis. In one study reported from Memorial Sloan-Kettering Cancer Center, multiple samples of the same specimen yielded different results in 25% of the cases.

How well this technique can be standardized for one institution may therefore be a problem. How well this technique can be standardized for one group is a major problem which must be addressed if analysis of the results of these breast cancer protocols is to be meaningful.

It is clearly essential that standardized pathology review mechanisms be established for all participating institutions for essentially all malignancies. While this is not the sole contribution pathology makes to group studies, it is of fundamental importance underlying the validity of the patient entry in each study.

## DIAGNOSTIC RADIOLOGY REVIEW

The second component of establishing the appropriateness of the diagnosis is to determine the extent of the disease, ie, staging. The Canadian-American Study on Hodgkin's Disease required a lymphangiogram as part of the workup of the patient. A central review of these diagnostic roentgenograms was established; I believe this was the first such mechanism in group studies. After early discrepancies, there was close agreement between the review panel and the institution's radiologists. Now, with increasing frequency, computerized axial tomography scans are being required in many studies of brain tumors. Their application in areas below the head may prove important in the future. Isotopic scans are routinely requested in many studies. It is essential that diagnostic radiologists be involved in establishing acceptable standards of imaging techniques for patient workups in defining the extent of disease.

## SURGERY STANDARDS

The recently published compendium on staging by the American Joint Commission of the American Cancer Society should be the basis of the uniform definitions of extent of disease. Surgical staging has become an accepted and important component of oncology over the last decade. For example, the staging laparotomy in Hodgkin's disease is now considered to be fairly routine. Yet in the new intergroup study of early Hodgkin's disease in childhood, a large number of the staging laparotomies of the initial cases entered were in variance with protocol requirements. The spleen may or may not have been removed. Periaortic node sampling was not always done on both sides of vertebral bodies. Pelvic node sampling was not always done. A method of review of surgical procedures is necessary, particularly as clinical trials turn with increasing frequency to solid tumors intra-abdominally. Standardization of the surgical participation in protocols is becoming increasingly important in stratifying patients and in establishing and executing therapeutic

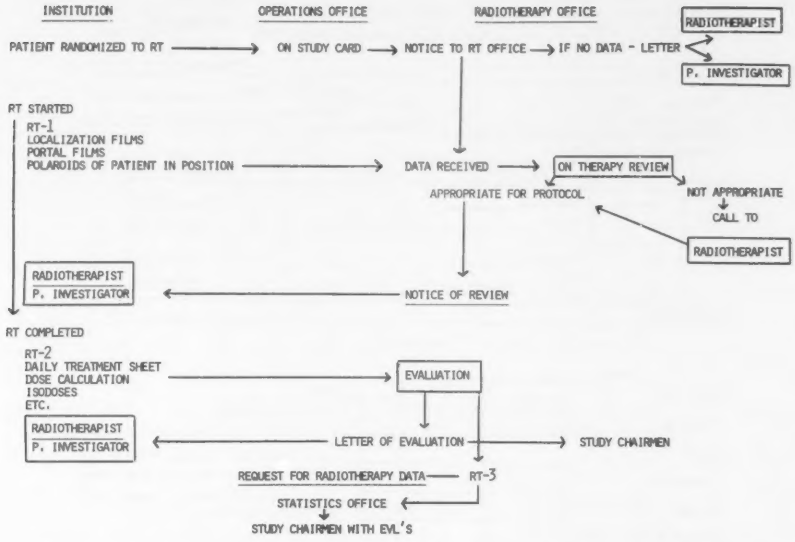



FIGURE 1.—CALGB Radiotherapy Committee Quality Assurance Program data flow diagram.

options. The review process to assure that the standards are observed must be established.

## RADIATION THERAPY REVIEW

With the development of multi-modality studies, the assessment of the appropriateness of treatment regimens delivered in each institution has undergone an expansion of complexity. It has become necessary to undertake a quality assurance program for each modality. In the CALGB the development of a Quality Assurance Program within the Radiotherapy Committee has been one of the major activities during the last 2 years. The program requires that data be submitted to the Quality Assurance Committee for all patients entered in radiotherapy arms of protocols within 3 days after the radiotherapy is started (fig 1). The purpose here is to assess the appropriateness of the treatment as required to fulfill protocol demands. If, on review, the Quality Assurance Committee detects a variation, the possibility exists that the treatment can still be modified at the institution so that it will no longer deviate from that required by the protocol. Accordingly, the number of patients who are appropriate for the study will increase, as will the evaluability rate.

For this evaluation, the radiotherapy treatment plan and radiographs to demonstrate the treatment volumes must be submitted for review by a radiotherapist and a physicist. Should it be determined that some difficulty exists with the treatment as planned, a telephone conversation with the responsible radiotherapist can modify the situation before the treatment has been completed. At the completion of treatment, a second form is submitted with all dose calculations as actually given and isodose curves that pertain. These are again reviewed and a final evaluation of the treatment is made. This assessment is forwarded to the study chairman and becomes a part of the permanent file on that patient's treatment. Deviations of 5%–10% are considered minor and $> 10\%$ are major protocol deviations (fig 2). The study chairman receives copies of all correspondence that has gone to the radiotherapists concerning the patient's treatment. He also receives a copy of the final evaluation form. When he is assessing the patient's response to treatment in the protocol, all background information and the appropriateness of the radiotherapy are available to him to complete his evaluation.

The Quality Assurance Program requires that there be a standardized review of each institution's physics program by the Radiological Physics Center (RPC), a group of physicists responsible to the American Association of Physicists in Medicine. The RPC determines the accuracy of the physical characteristics of the radiotherapy installation for all of the institutions that participate in group studies. Representatives of the RPC visit each institution to check the calibration of the machines used for the radiotherapy. In some instances, they may actually monitor a protocol by having the field arrangements simulated on a phantom.

The second aspect of our physics review requires that

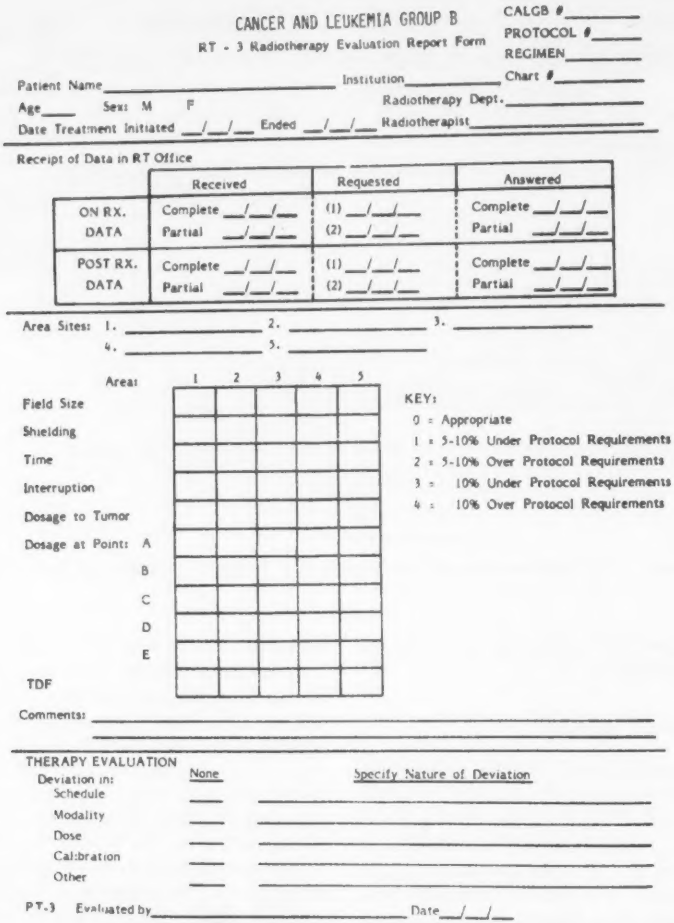

CANCER AND LEUKEMIA GROUP B
RT - 3 Radiotherapy Evaluation Report Form
CALGB #
PROTOCOL #
REGIMEN
Patient Name
Institution
Chart #
Age Sex: M F
Radiotherapy Dept.
Date Treatment Initiated __/__/__ Ended __/__/__ Radiotherapist
Receipt of Data in RT Office
| | Received | Requested | Answered |
|---|---|---|---|
| ON RX. DATA | Complete __/__/__ Partial __/__/__ | (1) __/__/__ (2) __/__/__ | Complete __/__/__ Partial __/__/__ |
| POST RX. DATA | Complete __/__/__ Partial __/__/__ | (1) __/__/__ (2) __/__/__ | Complete __/__/__ Partial __/__/__ |
Area Sites: 1. 2. 3. 4. 5.
Area: 1 2 3 4 5
Field Size
Shielding
Time
Interruption
Dosage to Tumor
Dosage at Point: A
B
C
D
E
TDF
KEY:
0 = Appropriate
1 = 5-10% Under Protocol Requirements
2 = 5-10% Over Protocol Requirements
3 = 10% Under Protocol Requirements
4 = 10% Over Protocol Requirements
Comments:
THERAPY EVALUATION
| Deviation in: | None | Specify Nature of Deviation |
|---|---|---|
| Schedule | | |
| Modality | | |
| Dose | | |
| Calibration | | |
| Other | | |
PT-3 Evaluated by Date __/__/__

FIGURE 2.—CALGB radiotherapy evaluation report form (RT-3).

isodose curves from each apparatus be filed with the Radiotherapy Office of the CALGB. Information concerning the method by which the dose distributions are calculated is included. Eighty-nine percent of all institutions in the CALGB either own their own computers or time-share. A computer program assists the CALGB physicist in the dosimetry review. The program is written in Fortran and runs on a DEC PDP 11/45 with remote entry terminals. Both the prescribed dose of radiotherapy and the calculations are reviewed at the start of treatment and upon completion of the prescribed course of treatment (fig 3).

With this, it is possible for the Quality Assurance Committee to review each institution's performance in terms of cases entered in radiotherapy arms of protocols, data received and evaluated, and appropriateness of the treatment. A score function has been produced based upon both the number of patients entered and the evaluable cases.

$$\left[1-\frac{\text{No data}}{\text{No. entered}}\right]\times\left[1-\frac{\text{Nonevaluable + major deviations}}{\text{Evaluated cases}}\right]\times 100 = \text{Performance score}$$

(a) (b) (c)

The first component in the equation (a) represents the percent of cases entered that actually can be assessed, and the second component expresses the evaluability rate of the cases for whom data have been submitted. Thus,

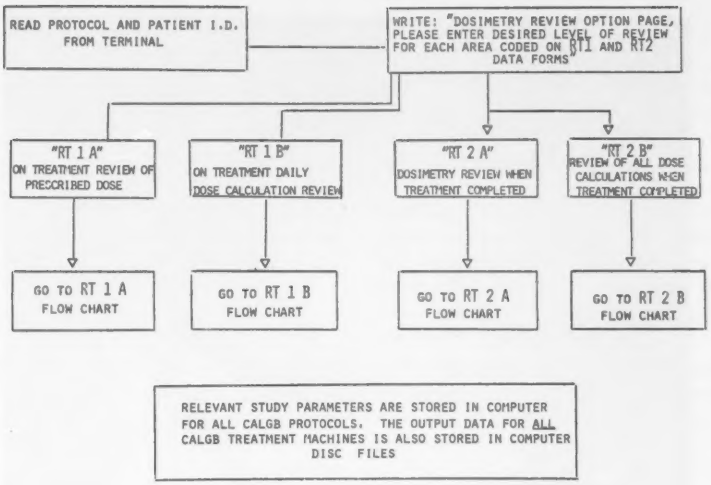

READ PROTOCOL AND PATIENT I.D. FROM TERMINAL
WRITE: "DOSIMETRY REVIEW OPTION PAGE, PLEASE ENTER DESIRED LEVEL OF REVIEW FOR EACH AREA CODED ON RT1 AND RT2 DATA FORMS"
"RT 1 A" ON TREATMENT REVIEW OF PRESCRIBED DOSE
"RT 1 B" ON TREATMENT DAILY DOSE CALCULATION REVIEW
"RT 2 A" DOSIMETRY REVIEW WHEN TREATMENT COMPLETED
"RT 2 B" REVIEW OF ALL DOSE CALCULATIONS WHEN TREATMENT COMPLETED
GO TO RT 1 A FLOW CHART
GO TO RT 1 B FLOW CHART
GO TO RT 2 A FLOW CHART
GO TO RT 2 B FLOW CHART
RELEVANT STUDY PARAMETERS ARE STORED IN COMPUTER FOR ALL CALGB PROTOCOLS. THE OUTPUT DATA FOR ALL CALGB TREATMENT MACHINES IS ALSO STORED IN COMPUTER DISC FILES

FIGURE 3.—Simplified flow chart for dosimetry review for CALGB.

if an institution enters 50 radiotherapy patients but submits data on only 45, the first expression would be 0.9. If, of the 45 patients, one was not evaluable because localization films were never taken and another one was considered a major deviation because the dose delivered was 12% low, the second expression (*b*) would be 0.96. The performance score (*c*) would then be 86. All three figures—the 90% compliance rate, the 96% appropriateness, and the performance score of 86—are made available to the Standards, Ethics and Peer Review Committee and to the Membership Committee of the Group. The score is made a permanent record of the institution's performance in the CALGB. Similarly, each protocol is reviewed in an attempt to determine how well the protocol is being followed, the problems that may exist with each study, and what can be done to clarify or alleviate some recurrent problems.

The application of this Quality Assurance Program has been well-received in the Group. Initially, just 1 year ago, only 20% of the required information of an assessment of patient treatment according to protocol actually was sent to the Radiotherapy Committee Office. Now, > 80% of this required information is being received and evaluated.

Some form of CNS prophylaxis has been part of CALGB protocols for acute lymphocytic leukemia for at least a decade (fig 4). The Radiotherapy Quality Assurance Program has been monitoring protocols CALGB-7611–7612 and found that a substantial number of radiotherapy centers were not conforming to the radiotherapy requirements (fig 5). We therefore decided to do a ret-

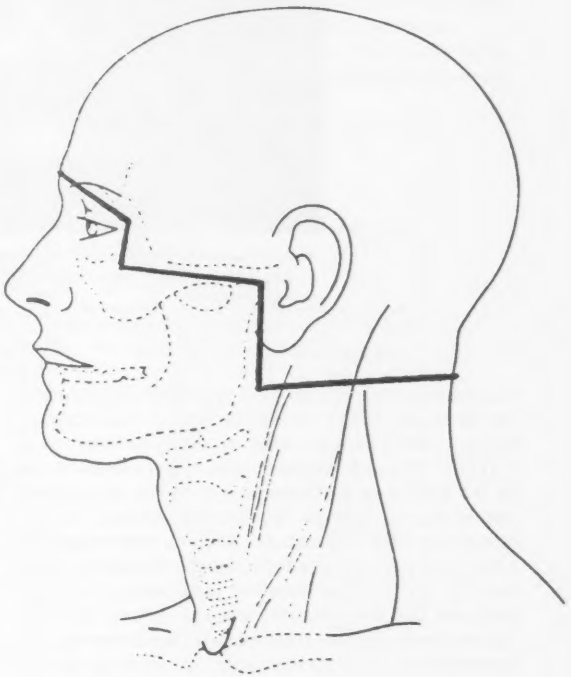


FIGURE 4.—Anatomic diagram for CNS prophylaxis.

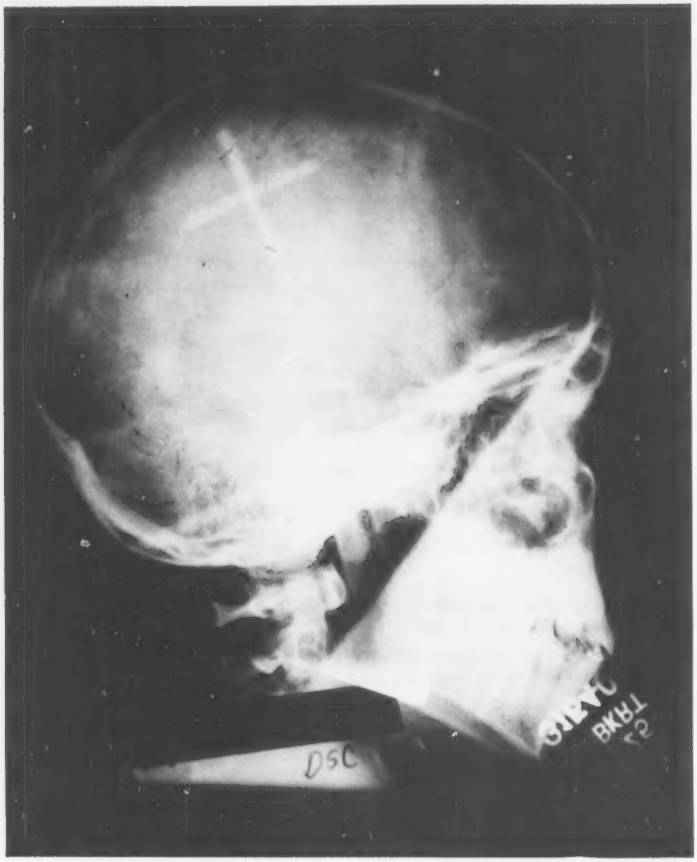


FIGURE 5.—Submitted localization film of CNS field inappropriately applied. Note that field does not include the entire base of the skull and excludes much of the posterior fossa.

rospective review of a protocol previously completed by the Group but before the formation of a Radiotherapy Review. Fifty patients were randomly selected from CALGB-7411, a predecessor protocol, and were analyzed by the Radiotherapy Committee as to the appropriateness of the radiotherapy field and the physics of the dosimetry and dose calculations. Records were traced for 45 of the 50 patients. The Radiotherapy Committee found that nine patients had received radiotherapy appropriately and that the dosimetry was also correct. In 36 instances the fields were inappropriate, the dosimetry was inappropriate, or some component of the data was not assessable. After this assessment, the Operations Office of the Group supplied data concerning the outcome of these cases. The nine "appropriate" cases had no CNS recurrences, but three of the other 36 cases had CNS failures. While not statistically significant, these results are sufficiently provocative to warrant a more extensive review of our past experience with CNS prophylaxis to determine the importance of the accurate delivery of the radiation and its relationship to outcome.

This is an example of a fairly simple standardized form of radiotherapy which was not appropriately given at major teaching hospitals and cancer centers or at smaller affiliated institutions. These same institutions which supplied the 45 cases for review participate in Eastern Cooperative Oncology Group, Radiotherapy Oncology Group, Gynecological Oncology Group, Children's Cancer Study

Group, Gastrointestinal Study Group, Brain Tumor Study Group, and National Surgical Adjuvant Study, to mention just a few.

Of some interest to us at this time is following "protocol deviates" in other protocols to determine how protocol violation affects outcome so that this aspect of treatment failure can be categorized. Thus, even though protocol deviation is not an acceptable performance pattern, these studies will make it possible to assess the data on "deviant" patients and thereby document another dimension of the therapeutic modality.

## MEDICAL ONCOLOGY MONITORING SYSTEM

For the chemotherapy component of protocols, monitoring of dose schedules depends upon the accurate recording of the information concerning the dose delivered, the toxicity, and the response to treatment. Drug dose does not lend itself to the same kind of retrospective review that can be done for radiotherapy or surgery. Careful investigators have maintained excellent flow sheets with substantial documentation of each phase of a study, but in a busy clinic it requires a dedication and fortitude which can on occasion escape the harried practitioner.

We have now readied for field testing a Medical Oncology Monitor System (MOMS) for chemotherapeutics at the Rhode Island Hospital Department of Radiation Oncology. As with its radiotherapy counterpart, the protocol monitor system is an interactive CRT-based program written in Fortran implemented on a DEC PDP 11/45. It makes use of a pagination system which provides the following information on the CRT (fig 6): patient's height and weight, temperature, performance index, toxicity index, hematologic status, and response at four different tumor sites. The data for three previous clinic visits are displayed. Data for this visit are entered on the terminal. The drug dose required for the patient's treatment according to protocol is then calculated for this visit and displayed on the monitor. Also displayed are the doses which the patient received on three previous visits as well as the cumulative doses over the entire course of the protocol (fig 7).

The attending physician retains the option to change a suggested dose. Should he desire to give a dose different from that displayed, he types in a dose modification. If this exceeds by 15% the protocol requirements, the physician is informed of a potential protocol violation. He can then (*a*) enter still another dose, (*b*) re-choose the computer-suggested dose, or (*c*) override the warning. If action (*c*) is taken, the patient's data file is flagged for future reference. Finally, the next visit page is displayed (fig 8). In addition to specifying the date on which the patient should return to the clinic, the system displays any laboratory tests required by the protocol on that date. Hard copy of this page is provided for clerical purposes.

The final component of the program currently being used is a protocol termination page which allows the entry

03-MAR-78

RHODE ISLAND HOSPITAL PROTOCOL MONITOR

*NAME: (LAST) XXXXX (FIRST) XXXXX (MI) __ *
*HOSPITAL # SR112457 PROTOCOL # CA7451 ARM # 3 *
*DATE OF RANDOMIZATION 10/30/76 DATE OF 'DAY 1' 03/18/77 *
*RESPONSIBLE PHYSICIAN XXXXXX XX, MD *

| PARAMETER | TODAY 08/12/77 | 08/05/77 | PREVIOUS VISITS 07/15/77 | 07/08/77 |
|---|---|---|---|---|
| HEIGHT (CM) | 163 | 163 | 163 | 163 |
| WEIGHT (KG) | 49.0 | 48.5 | 48.1 | 48.3 |
| TEMPERATURE (CENT) | 37.1 | 37.3 | 36.8 | 37.0 |
| KARNOFSKY INDEX | 100 | 100 | 100 | 100 |
| TOXICITY CODE | 0 | 0 | 0 | 0 |
| WBC | 6500 | 8500 | 6300 | 8250 |
| PLATELETS (X 1000) | 234 | 305 | 338 | 376 |
| HEMOGLOBIN (GRAMS %) | 10.9 | 10.9 | 11.2 | 10.9 |
| HEMATOCRIT | 34.0 | 33.0 | 38.1 | 32.8 |
| | | | | |
| RESPONSE AT SITE A | CR | CR | CR | CR |
| RESPONSE AT SITE B | CR | CR | CR | CR |
| RESPONSE AT SITE C | CR | CR | CR | CR |
| RESPONSE AT SITE D | CR | CR | CR | CR |

ENTER TODAYS PARAMETERS

FIGURE 6.—MOMS: CRT format of patient data.

```
03-MAR-78

**********    RIH PROTOCOL MONITOR    ----    DRUG DOSE DATA    **********
* NAME: (L) XXXXX                (F) XXXXXX                SURF. AREA 1.51 SQ.M  *
*************************************************************************************
*********************** TODAY ********** PREVIOUS VISITS *********CUMULATIVE*
*    DRUGS + DOSES     * 08/12/77  *08/05/77  *07/15/77  *07/08/77  *   DOSES  *
*************************************************************************************
*BCNU                  * 0.0  MG  *          *          *          *720.0  MG *
*(80.00 MG/M2 IV BOLUS)*  (100%)  *  (100%)  *  (100%)  *  (100%)  *  (100%)  *
*VINCRISTINE           * 0.5  MG  *          *          *          *7.6    MG *
*(1.400 MG/M2 IV BOLUS)*  (24 %)  *  (24 %)  *  (24 %)  *  (24 %)  *  (30 %)  *
*PROCARBAZINE          * 150.0 MG *          *          *          *1700.  MG *
*(100.0 MG/M2 ORALLY   )*  (100%)  *  (100%)  *  (100%)  *  (100%)  *  (100%)  *
*PREDNISONE            * 0.0  MG  *          *          *          *1680.0 MG *
*(40.00 MG/M2 ORALLY   )*  (100%)  *  (100%)  *  (100%)  *  (100%)  *  (100%)  *
*                      *_____ MG  *          *          *          *_____  MG *
*(_____ MG/__ _________)*  (___%)  *  (___%)  *  (___%)  *  (___%)  *  (___%)  *
*                      *_____ MG  *          *          *          *_____  MG *
*(_____ MG/__ _________)*  (___%)  *  (___%)  *  (___%)  *  (___%)  *  (___%)  *
*                      *_____ MG  *          *          *          *_____  MG *
*(_____ MG/__ _________)*  (___%)  *  (___%)  *  (___%)  *  (___%)  *  (___%)  *
*************************************************************************************
 OPTIONS:   1 = COMPUTER DOSES OK      2 = READ OTHER DOSE(S)
            3 = OVERRIDE PROTOCOL VIOLATION      ----------> _
ENTER OPTION
```

FIGURE 7.—MOMS: CRT format of drug dosage data.

```
03-MAR-78

*************************************************************************
*                  RHODE ISLAND HOSPITAL PROTOCOL MONITOR               *
*************************************************************************

              DATE OF NEXT REQUIRED VISIT -----> 09/08/77

            PROTOCOL REQUIRED STUDIES NEEDED AT NEXT VISIT
          ------------------------------------------------

                1.  CBC, PLATELETS
                2.  CHEST X-RAY
                3.  BUN, CREATININE
                4.  _____
                5.  _____
                6.  _____
                7.  _____
                8.  _____

          TYPE SHIFT/SEND TO RETURN TO OPTION PAGE
```

FIGURE 8.—MOMS: CRT format of followup page.

03-MAR-78

```
**********************************************************************
*                  RHODE ISLAND HOSPITAL PROTOCOL MONITOR            *
**********************************************************************
*    NAME: (LAST) XXXXX_____________ (FIRST) XXXXXX_____ (MI) _      *
*    PATIENT KEY SR112457__     PROTOCOL # CA7451   ARM # 3_         *
*    DATE OF RANDOMIZATION 10/20/77          DATE OF 'DAY 1' 03/18/77*
*    RESPONSIBLE PHYSICIAN XXXXXXX XX. MD_____________               *
**********************************************************************

REASON FOR PROTOCOL TERMINATION ----> 01      EFFECTIVE DATE 08/12/77

        1 = COMPLETION OF PROTOCOL TREATMENT
        2 = PROGRESSION OF DISEASE
        3 = EXCESSIVE TOXICITY
        4 = PATIENT REFUSED FURTHER TREATMENT
        5 = PATIENT LOST TO FOLLOW UP
        6 = DEATH OF PATIENT
        9 = OTHER: NOT SPECIFIED
```

FIGURE 9.—MOMS: CRT format of protocol termination page.

of seven reasons for the discontinuation of a study (fig 9). Hard copy of a tabulated up-to-date summary of the patient's clinic visits, key parameters, response data, and drugs administered will be provided.

The MOMS has been designed with a uniform reduction calculation for hematologic toxicity. Other toxic effects (neurologic, gastrointestinal, cardiac, etc) have also been accounted for, so that a uniform toxicity scale is displayed as part of the program. All parameters of toxicity can be assessed before determining the appropriate dose of chemotherapy to be given at this sitting. The CALGB has recently developed a uniform toxicity grading system for both radiotherapy and chemotherapy which is to be the basis of this part of the program.

Parenthetically, dose modification based upon a uniform toxicity grading is also being developed for the radiotherapists. As with chemotherapy, where the therapist can reduce the protocol dose or withhold it completely, so the radiotherapist can adapt his dose essentially for the same degrees of hematologic or gastrointestinal toxicity, for instance. Uniformity in dose modifications is a regular component of the protocol design. Assessing the appropriate use of the dose modifications is part of the Quality Assurance Program in both radiotherapy and chemotherapy.

The MOMS will make it possible for us to assess the appropriateness of the dose-reduction system used by the clinician in each case. It will then be possible to quantitate the implications of dose reduction by toxicity on outcome both in terms of ameliorating the toxicity and the course of the disease. Clearly, if a protocol is to be uniformly handled throughout the group, management of toxicity must be uniform. In evaluating the overall outcome of the study, Quality Assurance Programs must assess how well that is being done.

This MOMS may prove extremely valuable in insuring maximal protocol adherence and in encouraging specific positive action to modify treatment from protocol requirements on the basis of conscious selection rather than inadvertent oversight. The system is still new and being field-tested, and its ultimate place is yet to be determined.

# Use of New Biologic Agents in Clinical Immunotherapy[1]

**Brigid G. Leventhal**[2,*]

## SUMMARY

**The design of trials with biologic agents must consider the precise definition of the material (which may vary significantly from one laboratory to another), the formulation of the material (which is often complicated by its physical properties), and the safety of the material. To date, efficacy (defined as complete disappearance of tumor) has been seen only in situations associated with direct contact between tumor cells and adjuvant. The assessment of partial efficacy is most difficult because we know little about the basic changes which must be provoked in the immune system to achieve effective immunotherapy. It is only through careful assessment of the immune reaction of patients to their tumors and the correlation of these reactions with clinical response that we can hope to increase our knowledge.**

[Cancer Treat Rep 64:435–439, 1980]

New biologic agents in cancer are generally used in immunotherapy, although certain drugs such as L-asparaginase are biologic agents. Any new agent being considered for clinical trial must be a definable material. For most drugs this means a description of chemical structure, but this cannot be done for most biologic materials. The next requirement is for an adequate physical formulation; the material must be a standard product of known stability and must be available in a form suitable for administration. The next two requirements are those generally considered by the Food and Drug Administration, namely, safety and efficacy. These requirements are imposed with appropriate levels of stringency; a material which represents a new scientific concept, as do many immunotherapeutic materials, does not yet have standard guidelines. The best scientific judgment of the investigators and the regulatory agency is applied to each product to develop appropriate evaluation criteria.

Agents used for immunotherapy can be subdivided as can any other type of material used for immunization (1). First, immunotherapy can be specific, employing material which contains tumor antigens, or it can be nonspecific, being used to stimulate one or more components of the immune system in the hope that tumor cells will be killed incidentally as "innocent bystanders." Second, immunotherapy can be active, that is, inducing direct stimulation of the patients' own immune system, or passive, ie, administering to the patient the products of an immune reaction in another individual. Third, immunotherapy can be adoptive, involving the transfer to an individual of the "machinery" for continuous maintenance of specific immunity rather than a product of an immune reaction with a defined half-life (2).

The majority of agents for nonspecific immunotherapy will be used in the active mode. Such materials are often infectious agents or their products. These materials may be either live, such as the bacillus Calmette-Guérin (BCG) vaccine, or dead, such as *Corynebacterium parvum*. Partially purified materials such as the methanol-extracted residue (MER) of BCG (3) or a BCG cell-wall vaccine (4), which is often administered with a lipid emulsion, are also under investigation. Chemically defined materials such as levamisole are also used as nonspecific adjuvants.

As shown in figure 1, there are a number of sources of variability in the definition of the bacterial products which might be used in immunotherapy. First, the basic biologic variation in bacterial strains with time in culture can affect virulence or adjuvanticity; second, the method of culture can alter the characteristics of the bacteria. BCG itself was developed through a specific alteration

[1]Supported in part by Public Health Service grant CA-06973 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]Division of Pediatric Oncology, The Johns Hopkins Oncology Center, Baltimore, MD.

**Reprint requests to:* Dr. Brigid G. Leventhal, Division of Pediatric Oncology, The Johns Hopkins Oncology Center, 601 N Broadway, Baltimore, MD 21205.

DEFINITION
- Spontaneous (biologic) variability with generations in culture
- Variation with change in culture methods
- Different methods of inactivation

FORMULATION
- Transport form
- Reconstitution and administration may inactivate

SAFETY
- Potential infectivity to patient and patient contacts
- Potential virulence may parallel adjuvant effect
- Sensitivity to anti-infective therapy
- General inflammatory responses (fever, chills, abscess)
- Reactivation of previous infection (tuberculosis)
- Associated products or vehicles

FIGURE 1.—Problems associated with the use of infectious agents for active immunotherapy.

DEFINITION
- Chemical nature of product only partially defined
- Physical nature of product
  - Large particles, cell fragments, or aggregates
  - Degree of association may be critical

FORMULATION
- Emulsion or suspension not solution
- Sterilization without denaturation or disruption
- May require visual screen before administration
- Size may determine site of deposition

SAFETY
- Large particles--microemboli, secondary disseminated intravascular coagulation
- Inflammatory effects
  - Bronchospasm, disseminated granulomas, ulcers, fever

FIGURE 2.—Problems associated with the use of partially purified biologic materials in therapy.

in culture method; namely, the serial passage of *Mycobacterium bovis* in medium containing beef bile which produced an attenuation of virulence and made the material safe as a vaccine against tuberculosis (5). The significance of both of these factors is increased when one tries to compare strains of bacteria which have been carried for long periods in laboratories which are geographically separate from one another. It is not surprising therefore that the BCG from the Tice Laboratory in Chicago might have somewhat different biologic characteristics from the English Glaxo BCG (6). The same consideration would apply to other bacterial vaccines; for example, the *C. parvum* from the Pasteur Institute is likely to differ from the *C. parvum* prepared by the Burroughs Wellcome Laboratories. The method of inactivation of the bacteria for killed vaccines is another potential source of variability.

Variation in the material also occurs during formulation. The adjuvant effect of BCG clinically is related to the number of viable units of material administered (7). Intact BCG in lyophilized vaccines provokes a less vigorous immune response than material from log-phase cultures (6). In addition, the method of reconstitution in the recipient institution may affect the number of viable organisms. The method of administration may also be important, particularly when live material is given intradermally. Most procedures for skin preparation before injection involve cleansing with an antibacterial substance. One must demonstrate that the administered vaccine can survive the cleansing process or any locally applied drying methods. Ideally, these live materials should be initially studied in institutions which have the capacity to assess the viability of the product as it is administered to the patient rather than relying on the manufacturer's assessments.

The safety of these products is obviously related to their potential infectiveness, both for the patient and for his contacts in and out of the hospital. Risks which might be acceptable for a patient, such as the possible development of disseminated BCG disease, might not be acceptable for a pregnant woman or a newborn grandchild. The virulence of the organism must be monitored continually since it is always possible that a virulent substrain may re-establish itself.

Additionally, one must consider the potential sensitivity to therapy for an infection once established. No isoniazid-resistant line of BCG has ever been developed, but investigators who work with live BCG must consider this possibility, as well as the fact that these materials may provoke a generalized inflammatory response with fever, chills, and possible allergic phenomena. Reactivation of old tuberculous infections has been seen in patients overwhelmed by a large antigenic load of tuberculin.[3]

The final consideration in the administration of a vaccine is the effect of any associated products or vehicles which may be administered along with the organism. The principal difficulty with vaccines has been the administration of potentially allergenic culture materials. All of these factors also apply to bacterial systems and to viral or other infectious agents which might be used in immunotherapy.

As shown in figure 2, the use of partially purified products involves a somewhat different set of problems. The chemical nature of these products is only partially known and they usually contain residual undefined components. Often it is in fact the physical nature of the material which is most critical to its definition. Cell-wall components are administered adherent to oil microdroplets. If the droplets are of an inappropriate size or if they do not adhere in a specific pattern, the mixture will not exert its adjuvant effect (4). MER, which is principally protein, exists in aggregates of varying size (3).

The formulation of these materials for mass adminis-

[3]Bernstain I. CCSG Protocol No. 241, personal communication.

tration is extremely complicated. Neither emulsions nor suspensions can be prepared as reproducibly as solutions. Methods of sterilization must be employed which penetrate the irregular size of material adequately and yet do not denature the proteins upon which the biologic properties of the material depend. These materials may even require visual screening under a light microscope to ensure the proper physical condition of the material. The size of the aggregate is also critical to its delivery. Particles of varying sizes may filter into different areas of the reticuloendothelial system for removal, resulting in different patterns of immunostimulation.

Safety must also be considered when one is working with a partially purified biologic product. The particles, if of large size and administered iv, may act as microemboli and form secondary thrombus and possibly induce disseminated intravascular coagulation. Large particles deposited in the lungs may provoke an anti-inflammatory response which induces bronchospasm or diffuse granulomas. The associated inflammatory response to these materials may produce disseminated granulomas in many organs. Local ulceration at the injection site may be pronounced and long-lasting, and may flare up with each administration. This toxic side effect is likely to be more pronounced with killed antigen held in situ in adjuvant than the local response to a live bacterium which migrates to draining lymph nodes.

Before we can accept the concept of tumor-specific immunotherapy in man, we must accept the idea that tumor antigens exist in human systems. Active specific immunotherapy could employ whole tumor cells, cell fragments, or chemically defined antigens. If whole cells are being used (fig 3), their source must be considered, eg, from spontaneous tumor, autologous or allogeneic, or from cultured cell lines. The problems of definition in cultured cell lines are similar to those in cultured bacterial lines. Antigen may alter with prolonged time in culture or with passage through another living host such as a nude mouse, and antigen expression is affected by the phase of the cell in cycle. The antigenicity of cell lines is also affected by the method of inactivation, and there is general agreement that viable cells are stronger antigens than dead ones. The formulation of whole-cell populations for use in tumor immunotherapy poses a massive problem. On at least one occasion, however, a multi-institutional clinical trial has collected human tumor cells from several institutions, tissue-typed them, stored them in a viable state, and returned them to participating institutions for administration (7). It is encouraging to believe that if whole-cell vaccines were useful in cancer immunotherapy, means could be found to provide such cells on a nationwide basis.

When considering the safety of material that contains whole cells, the first concern is to determine whether the cells will grow in the recipient. This outcome is inversely proportional to the degree of antigenic difference between donor and recipient and is also related to the viability of the cell line and the immune status of the host. Another feature which must be considered is the accessibility for destruction of secondary tumor if growth does occur. The potential for reaction to nontumor antigens on the immunizing cell line must also be considered; these might be organ-specific and would provoke a degree of auto-immunity (8). With any specific manipulation of the immune system, the possibility of enhancement of tumor growth must be considered. Enhancement was originally thought to result when antitumor antibody coated antigenic sites on the tumor so that they were no longer available for attack by lymphocytes (9). More recently, it has been shown that enhanced tumor growth may also result from the stimulation of cellular immunity against tumor antigens (10,11). In general, small quantities of immunoreactive substances may enhance tumor growth while large quantities may be toxic or inhibitory; thus, almost any trial of immunotherapy may encounter a phase of enhanced tumor growth even if the eventual outcome of the trial is successful immune destruction of tumor. The problem of associated culture products is maximal when agents are adapted for growth in human cells. Human tissue culture cell lines should be screened for known infectious agents such as hepatitis virus and mycoplasma before they are used as vaccines. The possible presence of tumor viruses in this material remains a theoretic concern.

The problems associated with the use of cell fragments are essentially the same whether those fragments are bacterial or mammalian. It should be pointed out that circulating tumor antigen and antigen-antibody com-

DEFINITION
1. Source
   a. Autologous tumor
   b. Allogeneic tumor
   c. Tissue culture cell lines

FORMULATION
1. Collection method
2. Storage in viable state
3. Transport
4. Administration

SAFETY
1. Capacity for growth in recipient
   a. Antigenic differences
   b. Viability
   c. Immune status of host
2. Accessibility for destruction if growth occurs
3. "Contaminating" antigens, eg, organ-specific may provoke autoimmunity
4. Enhancement of tumor growth
5. Associated culture products
   a. Infectious agents: hepatitis, mycoplasma
   b. Preservatives
   c. ? Human oncogenic viruses

FIGURE 3.—Problems associated with the use of whole-cell vaccines for immunotherapy in man.

DEFINITION AND FORMULATION

1. Donor source
   a. Homologous
   b. Heterologous
2. Antigenic reactivity
   a. Tumor-specific
   b. Normal organ
   c. Other
3. Nature of reactivity
   a. Cytotoxic
   b. Cytophilic
4. Purification
   a. Serum
      Immunologobulin fraction
      Complement requirement
   b. Cells: reactive subtype

SAFETY

1. Foreign protein reactions
2. Reactions with normal organ antigens
3. Graft-versus-host disease
4. Enhancement of tumor growth

FIGURE 4.—Problems associated with the use of passive immunotherapy.

plexes are thought to be the "blocking factors" which interfere with antitumor cellular immune responses in vitro. Therefore, it is possible that a trial of the use of purified antigen for immunotherapy might well have an immunosuppressive phase.

The usual form of passive immunotherapy (fig 4) involves the use of antisera which may be homologous and are either organ- or tumor-specific. The effect of antisera on tumor cells will vary. Some will be directly cytotoxic and some will coat the cell and make it a target for attack by another component of the immune system. Methods for defining the specificity of antisera are better developed than those for the definition of cellular immune materials. Methods for purification of the various subclasses of immunoglobulins are also further advanced than cell-fractionation techniques. It is important, when attempting to use an antitumor serum, to define the complement requirements if such exist. Soon it may be practical to analyze and to better define cellular immune reagents and give an adequate description of purified cellular subtypes for human systems.

The dangers of passive immunotherapy include the classic forms of reaction to foreign protein including anaphylaxis and serum sickness. Reactions to normal organ antigens and enhancement of tumor growth may also occur. If the recipient is immunocompetent, after a certain time he may reject any cellular material that may be administered. If the recipient is immunosuppressed, however, and cellular material takes as a graft, he runs the risk of developing graft-versus-host disease which, in the case of bone marrow transplant, is lethal in about 25% of patients even if genetically matched siblings are used as donors (12,13).

Bone marrow transplantation could, in fact, be considered an example of nonspecific adoptive immunotherapy and has been reviewed recently (12,13). Transfer factor and immune RNA are other examples of materials which are being proposed as specific adoptive immunotherapy. There are problems in both the conceptual and chemical definition of these materials.

We come now to the most difficult part of our task, the assessment of efficacy for any or all of these antitumor reagents. It would be simple to say that a treatment is efficacious if tumor regresses completely and never recurs. When employed alone, however, most antitumor therapies are less than completely successful. How can we decide if our therapy is partially successful in man? The requirements for success with BCG immunotherapy in animals are: (*a*) the tumor must be small; (*b*) the host animal must have the ability to develop an immune response to mycobacterial antigens; (*c*) adequate viable organisms must be injected; and (*d*) the BCG and tumor cells must be in close proximity to one another. The ability of the host to develop an immune response to tumor-associated antigens may also be a factor (5). Even in animal systems, however, successful regression of a tumor nodule may not result in systemic tumor immunity (14). Certainly, evidence is accumulating in man that direct inoculation of tumor nodules with adjuvant will cause regression of the injected nodule (1). It is not clear that this will result in systemic immunity.

How would we assess systemic tumor immunity if it does not produce total regression of tumor? Since we do not know which immune mechanisms are required for tumor lysis, this is a difficult question. If we are to assess the efficacy of our antitumor therapy, we must evaluate the elements and subsets of the immune system which we have stimulated in the patients. First we must consider the route of administration of the immunogenic material. It is known that intradermal injection of an antigen may induce an immune response while iv injection of that same antigen may result in tolerance. On the other hand, Purnell and co-workers (15) have reported that in intact mice both systemic and intratumor injections of the nonspecific adjuvant *C. parvum* produced a reduction in tumor growth rate; however, if the mice were rendered athymic, the intratumoral injection of *C. parvum* was no longer effective yet the systemic injection still delayed tumor growth. This study suggests that the route of injection, even of nonspecific adjuvants, can determine the cellular nature of the immune reactivity to tumor. The relative timing of adjuvant and antigen injection is also important. With both protein (3) and lipid (16) adjuvants the cellular response to a tumor antigen may peak at different times after adjuvant injection than does the antibody response. Thus, one might maximize an immunization schedule to produce either a cellular or a humoral immune response to a particular tumor anti-

gen. The timing of immunization relative to chemotherapy is also critical. Many immune responses have a proliferative component. It might be possible to either abolish or increase the immune response by the timing of associated chemotherapy. The immune status of the host before immunotherapy must also be considered. It may be that a particular regimen will only work in a patient whose immune status is not overly depressed by tumor or therapy. On the other hand, several clinical trials have suggested that the beneficial effects of immunotherapy were limited to the immunosuppressed patients (17,18). All of these factors must be assessed in man after careful study of current knowledge so that we may begin to learn which aspects of the immune response are or are not important in a therapeutic antitumor response.

## REFERENCES


1. LEVENTHAL BG, and KONIOR GS. Immunologic treatment of neoplasms in man. *In* Mechanisms of Tumor Immunity (Green I, Cohen S, and McCluskey RT, eds). New York, NY, John Wiley and Sons, 1977, chap 9, pp 408–427.
2. MCKHANN CF, and GUNNARSSON A. Approaches to immunotherapy. Cancer 34:1521–1531, 1947.
3. WEISS DW, and YASHPHE DJ. Nonspecific stimulation of antimicrobial and antitumor resistance and of immunological responsiveness by the MER fraction of tubercle bacilli. *In* Dynamic Aspects of Host Parasite Relationships. New York, NY, Academic Press, 1973, vol 1, pp 133–223.
4. RIBI E, MILNER KC, GRANGER DL, ET AL. Immunotherapy with nonviable microbial components. Ann NY Acad Sci 277:228–238, 1976.
5. BAST RC, JR, ZBAR B, BORSOS T, ET AL. BCG and cancer (first of two parts). N Engl J Med 290:1413–1420, 1974.
6. MACKANESS GB, AUCLAIR DJ, and LAGRANGE PH. Immunopotentiation with BCG. I. Immune response to different strains and preparations. JNCI 51:1655–1667, 1973.
7. BAST RC, JR, ZBAR B, BORSOS T, ET AL. BCG and cancer (second of two parts). N Engl J Med 290:1458–1469, 1974.
8. BAECHTEL FS, and PRAGER MD. Anti-thymocyte autoimmunity in BALB/C mice accompanying immunization to a syngeneic lymphoma. J Immunol 118:175–179, 1977.
9. HELLSTROM KE, and HELLSTROM I. Immunological enhancement as studied by cell culture techniques. Ann Rev Microbiol 24:373–398, 1970.
10. PREHN RT. The immune reaction as a stimulator of tumor growth. Science 176:170–171, 1972.
11. FIDLER IJ. Inhibition of pulmonary metastasis by intravenous injection of specifically activated macrophages. Cancer Res 34:1074–1078, 1974.
12. THOMAS ED, STORB R, CLIFT RA, ET AL. Bone marrow transplantation (first of two parts). N Engl J Med 292:832–843, 1975.
13. ———. Bone marrow transplantation (second of two parts). N Engl J Med 292:895–902, 1975.
14. ZBAR B. Immunotherapy of bilateral lymph node metastasis by intralesional or paralesional injection of *M. bovis* (BCG). Proc Am Assoc Cancer Res and ASCO 19:87, 1978.
15. PURNELL DM, BARTLETT GL, KRIEDER JW, ET AL. Systemic (i.v. or i.p.) not intratumor (i.t.) treatment with *Corynebacterium parvum* (C.p.) inhibits mammary adenocarcinoma (CaD2) growth in thymectomized (Tx) antithymocyte serum (ATS) treated mice. Proc Am Assoc Cancer Res and ASCO 19:218, 1978.
16. PRAGER MD, and GORDON WC. Use of a lipid adjuvant in chemoimmunotherapy. Proc Am Assoc Cancer Res and ASCO 19:80, 1978.
17. CUTTNER J, GLIDEWELL OJ, and HOLLAND JF. A controlled trial of chemoimmunotherapy in acute myelocytic leukemia (AML). Proc Am Assoc Cancer Res and ASCO 19:333, 1978.
18. COHEN MH, CHRETIEN PB, IHDE DC, ET AL. Thymosin fraction V prolongs the survival of small cell lung cancer (SCLC) patients treated with intensive combination chemotherapy. Proc Am Assoc Cancer Res and ASCO 19:117, 1978.

# Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part I. Introduction [1]

**Arthur H. Schmale** [2,*]

## SUMMARY

**Psychosocial oncology is defined as a newly developing field which requires its own methodology for measuring the importance of psychologic and social factors for cancer patients, their families, and involved health professionals. As with all new areas of research, there must be a systematic approach once clinical observations and initial hypotheses have been made. The science for such an approach has been slow to develop because of the complexities of finding and testing the appropriate instruments for identifying and quantifying variables before trying to improve the patient's adjustment. In order to facilitate this process, a five-center collaborative group has been formed. Among its first tasks are the selection and testing of screening instruments.**

[Cancer Treat Rep 64:441–443, 1980]

Psychosocial oncology is a new or at least as yet poorly developed field devoted to the study of psychologic and social factors related to the cancer experience of patients, families, significant others, and health professionals providing care. The psychologic variables include knowledge, beliefs, attitudes, feelings, and behavior. The social factors are mainly those of personal, interpersonal, and institutional support. This field deals with those variables identified by some investigators as the "soft" sciences, and in time we plan to study the relationship of these variables to the biologic characteristics of the various forms of cancer and the reactions of these tumors or cells to the many forms of physical treatment. It should be noted that in psychosocial oncology we are looking not at cause-and-effect relationships but at covariate, facilitating, and inhibiting interreactions.

Cancer is greatly feared by the general public, it produces much morbidity and mortality, and the scientific community is exerting great efforts in its attempts to find effective diagnostic, therapeutic, and preventive measures. Thus, we are called upon to understand the public's concern, the patient's dilemma, and the professional's hopes and frustrations, and to determine how these forces interact; sometimes the interactions of these forces are to the patient's benefit, but more frequently they seem to reinforce feelings of uncertainty and in some cases produce unnecessary despair. In essence, psychosocial oncologists are attempting to understand the stresses associated with cancer and its treatments and the means by which these stresses are experienced and reacted to, ie, how these reactions facilitate or inhibit adequacy of adjustment, adherence to treatment, and restoration of equilibrium following treatment.

Those of us in the behavioral sciences who have had some experience working with cancer patients and have also studied psychopathology and psychosomatic relationships know that some of these assessment instruments may be useful in the study of cancer patients while others may not. Thus, we are convinced that it is necessary to start from scratch and build a general research approach which can be pursued in a systematic way. This approach will proceed in the following sequence.

*A.* Systematically identify the psychologic and social factors which promote or interfere with optimum patient participation and care (the selection or design of screening instruments and protocols to explore clinically identified problem areas).

*B.* How do these factors change for better or for worse during the progression from diagnosis to treatment and

[1]Supported in part by Public Health Service grants CA-19681 and CA-11198 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]Psychosocial Medicine Unit, University of Rochester Cancer Center, NY.

**Reprint requests to:* Arthur H. Schmale, MD, Department of Psychiatry, University of Rochester Medical Center, 300 Crittenden Blvd, Rochester, NY 14642.

then through the long-term followup (protocols with quantitative methods for measuring amounts and predicting change)?

*C*. How often do these problems occur? Are they of sufficient frequency and magnitude to justify further systematic study (sampling studies designed to be implemented by other psychosocial investigative units or treatment trial groups)?

*D*. What means are available to maximize adjustment or to reduce or prevent inadequate adjustment (selection or design of intervention procedures for their comparative effectiveness in improving adjustment)?

*E*. By what means can the relevant knowledge and skills be translated into optimum care and outcome? What roles, information, and responsibilities should the patients, families, and health professionals have in identifying, assessing, and intervening in these aspects of cancer care (publications, workshops, seminars, training, and education programs)?

*F*. Identify final measures of effectiveness (patient and family satisfaction, attitudes, survival time, practices, visits, and costs).

Currently, we in the psychosocial oncology field see ourselves where the cancer treatment field was 15 years ago in its development. With a systematic and collaborative approach we do not think it will take us 15 years to catch up. Why are we only now becoming concerned about psychosocial research? We know that over the past 30 years public policy has been greatly influenced by economic indicators. Material and monetary needs were considered to be the complete or dominant measure of the quality of national life and sense of well-being. Campbell (1), among others, thinks we are moving in the direction where the goals of our society are becoming more psychologic and introspective rather than just economic and demographic. How does all of this influence the cancer patient, family, and health professional? I think it is becoming obvious that scientific study, public education, and mass communication have made for a better informed public who have a greater awareness of all the uncertainties associated with the diagnosis and treatment of cancer. The public and the cancer patient want to know how they are going to be affected and want some say in the decisions that may determine the quality as well as the quantity of their lives. This need for a sense of control, belonging, achievement, and purpose helps to explain the interest in and support for psychosocial research.

Where are we in our methodologic approaches in this new field of psychosocial oncology? We have a number of isolated bits and pieces which represent what we are coming to call phase I, or clinical inquiry research as Tukey (2) refers to it; this is the first level on the approach list. Much of this research is based on clinical and interview observations that have been described in operational terms as a basis for very crude hypothesis testing. For the most part, the instruments for identifying the psychosocial data have been borrowed from empirically derived personality inventories, adjective feeling or mood check lists, and measures of gross psychopathology. In general, findings or results have correlated personality and life-event data to delay behavior in seeking diagnosis, personality characteristics related to already known specific tumor types, and survival time. Each investigation and investigator has varied the approach so that there has been little or no replication or building of one set of findings on another. In part, this can be attributed to the way in which funds are awarded in a new area where the relevance of the findings has yet to be valued. Each piece of research has to be constructed as more or less original and with the intent of establishing new and, it is hoped, more important findings. The time allotted and the money provided never seem enough to go beyond the exploratory study stage.

Searching for a means of breaking through this log jam, several psychosocial investigators approached the National Cancer Institute (Bethesda, MD) and the Division of Cancer Control and Rehabilitation of that institute seeking support for a small collaborative group to study the field in a more systematic way. Five investigative units and a resource center for data handling were funded to participate in this venture. Since we are only in the planning phase of this research, I can only tell you of our general approach and what we see as the areas which have a reasonable hope of systematic investigative success in our hands over the next several years. These areas are limited to the treatment and followup periods since it is in these phases that the patient is most available to us for study. They include (*a*) what the patient knows and how he reacts to what he has been told with regard to diagnosis and treatment, (*b*) reaction to the informed-consent procedure at several points in treatment time, (*c*) reaction to treatment delay and side effects (especially nausea and vomiting), (*d*) psychotropic drugs and their effects, (*e*) reaction to recurrences or metastases, and (*f*) reaction at the end of and after treatment.

The five-center group has identified three types of instruments it wants to use for psychosocial screening (the identification of psychosocial problems). These three measures are psychosocial adjustment or coping, social supports, and perceived well-being. These measures will be crude at first, but they are important for our systematic identification of problem areas or subgroups. We also are not sure how distinct and separate these three sets of variables are. In essence, they probably represent three different but related areas.

An important distinction can be made between where psychosocial oncology is at the present time (the stage of identifying psychosocial strengths and problems and agreeing on how to recognize them) and where the physical and biological sciences are (conducting clinical trials to determine the magnitude or amount of change and the

qualitative characteristics of the change associated with already identified disease and established treatment characteristics).

In our measures of psychologic adjustment, social support, and general well-being, we are concerned about how we are to compare the so-called subjective measures of behavior with the biologic variables of bodily functioning which are thought of as being more objective and accurate as measures of state of health, disease, and general well-being. Many investigators still perceive the patient's evaluation of how he or she is doing (based on feelings, attitudes, and functioning) as subjective data and therefore potentially unreliable, impossible to interpret, and too incomplete to be of value. Andrews (3), at the University of Michigan School of Social Research, makes the point that the objective measures which involve the counting of occurrences of given physical, psychologic, and sociodemographic phenomena are only indirect measures of well-being and that the subjective or more direct measures can be developed, measured, monitored over time, analyzed, and interpreted. Andrews notes that, in spite of improving material and economic conditions, the blacks in the US during the early 1960s experienced a simultaneous increase in manifestations of discomfort, and he makes the following point: "there is reason to believe that there exists only a loose linkage between the objective conditions of people's lives and the individual perceptions of well-being."

Many investigators working in oncology drug treatment trials use the Karnofsky performance status rating as a general indicator of physical adjustment and consider this adequate to measure and follow the patient's condition (4). The performance status criteria classify the patient's level of activity and functional capacity from 100% to 0. The ratings range from 100% to 80% where the patient is able to carry on normal activity with little or no effort, symptoms, or evidence of disease; from 70% to 50% where the patient is unable to work but able to live at home requiring more or less assistance; and from 40% to 0 where the patient is unable to care for self, requires the equivalent of institutional care, and has rapidly progressing disease and becomes moribund, and is rated 0 at the time of death. This instrument has proved useful in the drug treatment trial collaborative protocols and correlates with disease and physical state. Frequently, however, it does not correlate with the patient's perceptions and feelings of how things are going. (Some investigators even suggest that part of the unaccounted-for variance in the patient's response or lack of response to a specific treatment trial resides in the patient's attitude or beliefs.) Several studies have now reported little or no correlation between physical and psychosocial adjustment or an inverse relationship between these two parameters at the time of followup after treatment.

Our psychosocial oncology unit has just completed a study on the psychosocial adjustment of cancer patients who had received radiation therapy. There was a significant correlation between the psychosocial difficulties reported 1 month into treatment and the psychosocial adjustment reported at 6 months for 44 patients and again at 18–24 months for 24 patients after treatment began.[3] Interview data were scored on a rating system of 0–2 for each of three categories of psychosocial adjustment, namely, intrapsychic, social, and abnormal illness behavior. In contrast to the significant association between these psychosocial problem scores over time, there was an inverse relationship between the Karnofsky performance status and the psychosocial problem scores at both 6 and 24 months. These findings are consistent with those of Eisenberg and Goldenberg (5), who studied 252 patients at 1 and 18 months after mastectomy, and with the report of Craig et al (6) on the quality of 5-year survival in 134 patients with breast cancer and 260 normal controls.

Other issues of importance include the recording of perceptions by patients versus other observers. We recognize the importance of having self-rating as well as observer-rating forms of the same scales. One would not be more accurate or objective than the other, but together they should help us first to recognize the magnitude of the differences or the degree of similarity and second to study the reasons for the similarities and the differences. As indicated earlier, we also think that the health professional's actions and the patient's family members or significant others' attitudes, beliefs, and actions are important variables to measure in our attempts to understand the cancer patient's psychologic adjustment and his/her ultimate sense of well-being.

## REFERENCES


1. Campbell A. Subjective measures of well-being. Am Psychol 31:117–124, 1976.
2. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. Science 198:679–684, 1977.
3. Andrews FM. Social indicators of perceived life quality. Soc Indicators Res 1:279–299, 1974.
4. Karnofsky DA. Cancer chemotherapy. Clinical evaluation of anticancer drugs. Gan Monogr 2:223–234, 1974.
5. Eisenberg HS, and Goldenberg IS. A measurement of quality of survival of breast cancer patients. *In* Clinical Evaluation in Breast Cancer (Hayward JL, and Balbrook RD, eds). New York, NY, and London, England, Academic Press, 1966, pp 93–107.
6. Craig TJ, Comstock GW, and Geiser PB. The quality of survival in breast cancer: a case-control comparison. Cancer 33:1451–1457, 1974.


[3]Schmale AH, Morrow GR, Davis A, et al. Psychosocial adjustment of cancer patients. Presented at the annual meeting of the American Psychosomatic Society, Washington, DC, Mar 31, 1978.

# Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part II. Assessing Reliability With Equal-Appearing Interval Scales [1]

**Michael L. Feldstein** [2,*]

## SUMMARY

**The usual statistical procedures for assessing reliability among several raters do not apply to scales of the equal-appearing interval type. This is so primarily because the set of possible responses on such scales, namely the nonnegative integers 0,1,. . . (M-1), contains so few points that normal distribution theory cannot be invoked. Traditional analyses of such data, which include the one-way analysis of variance intraclass correlation, produce a reliability coefficient based on an assumed normally distributed response variable, and must therefore be interpreted with great caution. This paper discusses a statistical approach which exploits the discrete nature of the response variable and hence is more appropriate for such data.**

[Cancer Treat Rep 64:445–450, 1980]

Equal-appearing interval (E-1) scales are those in which the responses can be quantitatively labeled by the nonnegative integers 0,1,2,. . . (M-1) for a specified value of M. That is, the label j, j = 0,1,. . . (M-1) refers to a score of magnitude j, and not to a categoric response labeled j. The statistical analysis of data of the latter type has been discussed recently in a series of papers by Landis and Koch (1–3), and earlier by Fleiss (4) and Light (5).

The fact that E-I scales are quantitative has led many authors to treat the response variable as though it were normally distributed. Thus, Winer (6) proposes that the data be analyzed like a random effects model (model II, one-way analysis of variance) which produces a reliability coefficient which is a simple function of the two error variances. This approach can occasionally lead to rather embarrassing results as when one of the error variances has a negative estimate. In such cases, instead of questioning the model, the reliability is usually reported as zero. Bartko (7), after reviewing the most frequently used reliability measures including the Spearman-Brown prediction formula and the Winer "anchor-point" model, argues against the use of many of them. His preference seems to be for the one-way analysis of variance intraclass correlation (8) which produces a high-reliability coefficient only if (unlike the other methods he discusses) the within-items mean square is small relative to the between-items mean square.

The assumption that scores on an E-I scale are normally distributed *may* be tenable if M is sufficiently large, but in most practical situations it is not. Typically, M is $< 12$, as in the Karnofsky scale and in its abbreviated version currently used by many national cancer clinical trial groups. Parenthetically, it is worth asking whether the Karnofsky rating scale, or indeed any such "apparent" E-I scale, is actually an E-I scale or rather, a more primitive ordinal one. In any event, if such data are to be treated quantitatively, we must abandon the usual procedures of Bartko (7,8) and others wherein the response variable (score) is assumed to be continuous.

To this end, we imagine an E-I scale with M values 0,1,. . .,M-1, being utilized by K raters, independently judging n items. We shall consider only the "balanced" case so that the data can be displayed as in table 1. The typical entry in table 1, namely $l_{ij}$, can be regarded as the outcome of a discrete random variable $L_{ij}$, which takes values in the set $S_M = \{0,1,2,. . .,M-1\}$ and refers to the score rater j assigns to item i. We shall assume not only

[1]Supported in part by Public Health Service grant CA-11198 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]Division of Biostatistics, University of Rochester, NY.

**Reprint requests to:* Michael L. Feldstein, PhD, Division of Biostatistics and Epidemiology, Sidney Farber Cancer Institute, 44 Binney St, Boston, MA 02115.

TABLE 1.—Typical data display for K raters judging n items

| Item being judged | Rater | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | . | . | . | K |
| 1 | $l_{11}$ | $l_{12}$ | $l_{13}$ | . | . | . | $l_{1K}$ |
| 2 | $l_{21}$ | $l_{22}$ | $l_{23}$ | . | . | . | $l_{2K}$ |
| 3 | $l_{31}$ | $l_{32}$ | $l_{33}$ | . | . | . | $l_{3K}$ |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| n | $l_{n1}$ | $l_{n2}$ | $l_{n3}$ | . | . | . | $l_{nK}$ |

that raters judge independently, but also that for any rater, items are judged independently. This situation may not always obtain, especially if there is a "learning" effect by the rater as the items are judged serially. We let

$\Pr(L_{ij}=l_{ij}) = p_{ij}(l_{ij})$, $i=1,2,\ldots,n; j=1,2,\ldots,K; l_{ij}\varepsilon S_M$. Clearly,

$$\sum_{l=0}^{M-1} p_{ij}(l) = 1$$

and we suppose that the probability that item i receives score $l$ from judge j depends generally both on the item (the stimulus) and the rater. The joint distribution of the nK random variables $l_{ij}$ is just the product

$$\prod_{i=1}^{n}\prod_{j=1}^{K} p_{ij}(l_{ij})$$

and the hypothesis of interest is

$$H_0\colon\ p_{ij}(l) = p_i(l) \text{ for every j, and any i.}$$

$H_0$ is equivalent to declaring that the raters agree with one another. We propose to test $H_0$ conditional upon the total scores received by each item; ie, the n row sums of table 1. If we let

$$\sum_{j=1}^{K} l_{ij} = l_i, \qquad i=1,2,\ldots,n$$

then the required conditional joint distribution or conditional likelihood is given by

$$L^C = \prod_{i=1}^{n}\left\{\frac{\prod_{j=1}^{K-1} p_{ij}(l_{ij})p_{iK}(l_i-l_{i1}-\cdots-l_{i,K-1})}{\sum' \prod_{j=1}^{K-1} p_{ij}(x_{ij})p_{iK}(l_i-x_{i1}-\cdots-x_{i,K-1})}\right\} \quad [1]$$

where $\Sigma'$ denotes summation over the K-1 variables $x_{i1},x_{i2},\ldots,x_{i,K-1}$ such that

$$x_{ij}\varepsilon S_M,\ j=1,2,\ldots,K-1, \text{ and } \sum_{j=1}^{K-1} x_{ij} \le l_i,\ i=1,2,\ldots,n.$$

Equation [1] is the starting point for inference regarding the hypothesis $H_0$. To proceed further, we must make some assumptions regarding the sequence of probabilities $p_{ij}(l)$, $l=0,1,2,\ldots,M-1$. In the next section, we shall examine some possible models for $p_{ij}(l)$, all of which, in the absence of replication, cannot be tested. While this is a severe shortcoming, the same criticism applies to the traditional approaches taken by Bartko (8) and others, who also assume a normal distribution for the response variable, ie, the scores. We remark that in most "reliability" studies, replication is generally not possible.

## PROBABILITY MODELS FOR THE RESPONSE VARIABLE

The Poisson distribution is commonly assumed to be a model for independent counts in a contingency table [eg, Bishop et al (9)]. If we restrict the Poisson random variable to the set $S_M$, then we have the model

$$p_{ij}(l) = \frac{\mu_{ij}^{l}/l!}{\sum_{r=0}^{M-1} \mu_{ij}^{r}/r!} \qquad l=0,1,\ldots,M-1$$

where $\mu_{ij}$ is an unknown parameter which reflects both on the item being judged and the rater doing the judging. We shall make the further structural assumption that

$$\mu_{ij} = \theta_i\phi_j,\ i=1,2,\ldots,n;\ j=1,2,\ldots,K$$

where $\theta_i$ is the effect due to the $i^{th}$ item and $\phi_j$ is the effect due to the $j^{th}$ rater. This choice of model, which is clearly a log-linear one with no interaction term (9), has the following physical interpretation. Suppose that the true score for item i is a Poisson random variable with mean

$\theta_i$ and that each rater views or judges each item for exactly the same duration. Within this fixed time frame we further suppose that each rater never imagines an occurrence of the event being counted. Let $\phi_j$ be the constant probability of detecting the event by the jth rater. That is, we also assume no learning takes place during the course of rating the items which would alter the detection probability. Then it is easily shown that for item i the probability that rater j detects exactly X occurrences follows the Poisson distribution with mean $\mu_{ij} = \theta_i\phi_j$. In terms of the original hypothesis $H_0$, the n quantities $\theta_1$, $\theta_2,\ldots,\theta_n$ are nuisance parameters. It is straightforward to show that conditioning on the row sums of table 1 completely removes the nuisance parameters, and we get

$$L^c = \prod_{i=1}^{n} \left\{ \frac{\prod_{j=1}^{K-1} \left(\frac{\phi_j}{\phi_K}\right)^{x_{ij}} \Big/ \prod_{j=1}^{K} x_{ij}!}{\sum' \prod_{j=1}^{K-1} \left(\frac{\phi_j}{\phi_K}\right)^{x_{ij}} \Big/ \prod_{j=1}^{K} x_{ij}!} \right\} \qquad [2]$$

where $l_{iK} = l_i - l_{i1} - l_{i2} - \ldots - l_{i,K-1}$

and $x_{iK} = l_i - x_{i1} - x_{i2} - \ldots - x_{i,K-1}$.

Under the multiplicative (independent) structural assumption, the hypothesis $H_0$ is equivalent to the hypothesis

$$H_{00}: \phi_1 = \phi_2 = \ldots = \phi_K$$

for then $p_{ij}(l)$ does not depend on the rater effect $\phi_j$. The conditional likelihood function given by equation [2] can be further simplified if we let M tend to infinity. In that case, the set of scores includes all nonnegative integers, and the score becomes an unrestricted Poisson random variable. In view of the conditioning, $L^c$ reduces to the product of n independent multinomial distributions, the $i^{th}$ being

$$\text{Mult}[l_i;\ a_1/A, a_2/A, \ldots, a_K/A]$$

where $a_j = \phi_j/\phi_K$, $j = 1,2,\ldots,K$; and $A = \sum_{j=1}^{K} a_j$

Furthermore, under $H_{00}$, the conditional likelihood $L^c$ corresponds to n independent multinomial distributions, the $i^{th}$ being

$$\text{Mult}[l_i;\ 1/K, 1/K, \ldots, 1/K].$$

Two asymptotically equivalent (K-1 degrees of freedom) test statistics are immediately available for testing $H_{00}$, the likelihood ratio test statistic, and the Pearson goodness-of-fit statistic. The former is given by

$$\Lambda = K^{-N} \Big/ \prod_{j=1}^{K} (c_j/N)^{c_j} \qquad [3]$$

and the latter by

$$X^2 = \sum_{j=1}^{K} \frac{(c_j - N/K)^2}{N/K} \qquad [4]$$

where $c_j = \sum_{i=1}^{n} x_{ij}, \quad j=1,2,\ldots,K;$

and $N = \sum_{i=1}^{n} l_i = \sum_{j=1}^{K} c_j$

Both test procedures depend only on the column (rater) sums and not at all on the row (item) sums. They therefore lead to inferences which are item-independent, in sharp contrast to the approaches discussed by Bartko (7), all of which depend crucially on the "between items" mean square. Asymptotically, $-2\log_e\Lambda$ and $X^2$ are chi-square distributed each with K-1 degrees of freedom if the null hypothesis is true. These two equivalent test statistics are simple to use but rely on the assumption that the unrestricted Poisson variable is a good approximation to the one truncated to the set $S_M$. We point out, however, that the models presented here cannot distinguish between complete agreement and complete disagreement. The latter can occur when two raters score in precisely opposite directions. The necessary refinement to accomodate such a situation is the subject of another paper. Some insight regarding the appropriateness of this assumption is provided by examining the conditional likelihood function (equation [2]) when K=2. In this case,

$$L^c = \prod_{i=1}^{n} \left\{ \frac{\binom{l_i}{x_{i1}} \left(\frac{\phi_1}{\phi_2}\right)^{x_{i1}}}{\sum_{x_{i1}} \binom{l_i}{x_{i1}} \left(\frac{\phi_1}{\phi_2}\right)^{x_{i1}}} \right\}$$

and clearly for fixed i

$$\sum_{x_{i1}} \binom{l_i}{x_{i1}} \left(\frac{\phi_1}{\phi_2}\right)^{x_{i1}} = \sum_{x_{i1}=0}^{\min(l_i, M-1)} \binom{l_i}{x_{i1}} \left(\frac{\phi_1}{\phi_2}\right)^{x_{i1}}$$

Since we are conditioning on the row totals, $l_1,l_2,\ldots,l_n$, the quantities $\min(l_i,M-1)$, $i=1,2,\ldots,n$ are fixed numbers. Thus, if we re-order the rows of table 1 so that the first r rows correspond to items such that

$$\min(l_i,M-1) = l_i,$$

while for the remaining n-r rows

$$\min(l_i,M-1) = M-1,$$

we may write

$$L^C = \frac{\left(\frac{\phi_1}{\phi_2}\right)^{c_1} \prod_{i=1}^{n} \binom{t_i}{t_{i1}}}{(1 + \phi_1/\phi_2)^N \times D} \qquad [5]$$

where

$$D = \prod_{i=r+1}^{n} \left\{ 1 - \frac{\binom{t_i}{M}\left(\frac{\phi_1}{\phi_2}\right)^M + \binom{t_i}{M+1}\left(\frac{\phi_1}{\phi_2}\right)^{M+1} + \ldots + \binom{t_i}{t_i}\left(\frac{\phi_1}{\phi_2}\right)^{t_i}}{\left(1 + \frac{\phi_1}{\phi_2}\right)^{t_i}} \right\}$$

If r=n, then D=1, which would also be the value of D if the response variable were unrestricted Poisson In either of these cases the appropriate test statistic is given by equation [3]. Generally $0 \leq r < n$, however, and D will be positive and something less than unity. In this case, a closed form expression for the likelihood ratio test statistic cannot be obtained, although for given data, $I^c$ given by equation [5] can be maximized numerically by choice of $\phi_1/\phi_2$ (a first approximation to which is $c_1/c_2$). Thus the test statistic can be computed numerically and compared to the one given by equation [3]. We remark that, unlike the test statistic for the Poisson model, the test based on equation [5] for K=2, and on equation [2] for arbitrary K>2, does indeed depend on the row (item) sums.

For values of K larger than two, the situation is considerably more complex, at least theoretically. This suggests that we should use the test statistic given by equation [3] or alter the underlying probability model. We discuss one such alternative model below.

In this section we examine the hypothesis of rater agreement based on the conditional likelihood given by equation [1], when the underlying probability model corresponds to the truncated geometric distribution. That is,

$$Pr\left\{L_{ij} = t\right\} = \mu_{ij}^{t} \Big/ \sum_{r=0}^{M-1} \mu_{ij}^{r} \qquad t=0,1,\ldots,M-1.$$

Again, we make the structural assumption

$$\mu_{ij} = \theta_i \phi_j$$

and analogous to equation [2], we have for the conditional likelihood

$$L^C = \prod_{i=1}^{n} \left\{ \frac{\left(\frac{\phi_1}{\phi_K}\right)^{t_{i1}} \left(\frac{\phi_2}{\phi_K}\right)^{t_{i2}} \ldots \left(\frac{\phi_{K-1}}{\phi_K}\right)^{t_{i,K-1}}}{\sum' \left(\frac{\phi_1}{\phi_K}\right)^{x_{i1}} \left(\frac{\phi_2}{\phi_K}\right)^{x_{i2}} \ldots \left(\frac{\phi_{K-1}}{\phi_K}\right)^{x_{i,K-1}}} \right\} \qquad [6]$$

where $\Sigma'$ denotes summation over $x_{ij}$ such that $x_{ij} \varepsilon S_M$ and

$$\sum_{j=1}^{K-1} x_{ij} \leq t_i .$$

Once again, the n statistics $l_1, l_2, \ldots, l_n$ are sufficient for the n nuisance parameters $\phi_1, \phi_2, \ldots, \phi_n$. For K>2 this model is as unwieldy as the truncated Poisson model, but for K=2 the conditional likelihood simplifies considerably. In particular, when K=2, equation [6] becomes

$$L^C = \frac{\tau_1^{c_1}}{\prod_{i=1}^{n} \left\{ \sum_{r=0}^{z_i} \tau_1^{r} \right\}}$$

where $z_i = \min(l_i, M-1)$, i=1,2,. . .,n and $\tau_1 = \phi_1/\phi_2$. Thus

$$L^C = \frac{\tau_1^{c_1} (1-\tau_1)^n}{\prod_{i=1}^{n} (1-\tau_1^{1+z_i})} \qquad [7]$$

which is easily maximized numerically by choice of $\tau_1$. Denoting by $\hat{\tau}_1$ the maximum likelihood estimator of $\tau_1$, we have for the likelihood ratio test statistic

$$\Lambda = \frac{\prod_{i=1}^{n} (1+z_i)^{-1}}{L^C(\hat{\tau}_1)} \qquad [8]$$

As before, $-2\log_e \Lambda$ is asymptotically chi-square distributed with one degree of freedom if the null hypothesis is true.

Other models may be worth considering. However, in the absence of replicable reliability studies, the selection of a probability model remains an open question. In the next section we present some examples which suggest that the choice of a model may not be crucial.

## NUMERICAL EXAMPLES

### Two Rater Problems

Table 2 is a slight modification of three data sets reported by Bartko (7) in which the rating scale consisted of the integers 1 through 10 inclusive. To conform to the methodology outlined in this paper, we have reduced each score by unity so as to produce an E-I scale running from 0 to 9.

TABLE 2.—3 sets of ratings on a 0–9 scale (reprinted from ref 7)

| | Set 1 | | Set 2 | | Set 3 | |
|---|---|---|---|---|---|---|
| | Rater | | Rater | | Rater | |
| Item | 1 | 2 | 1 | 2 | 1 | 2 |
| 1 | 0 | 0 | 0 | 4 | 0 | 1 |
| 2 | 1 | 1 | 1 | 5 | 1 | 3 |
| 3 | 2 | 2 | 2 | 6 | 2 | 5 |
| 4 | 3 | 3 | 3 | 7 | 3 | 7 |
| 5 | 4 | 4 | 4 | 8 | 4 | 9 |

TABLE 3.—A comparison of the probability models for the data in table 2

| | Set | | |
|---|---|---|---|
| Model and test statistic* | 1 | 2 | 3 |
| Poisson: Pearson goodness-of-fit | 0.00 | 10.00 | 6.43 |
| Poisson; likelihood ratio, $-2\log\Lambda$ | 0.00 | 10.47 | 6.64 |
| Truncated Poisson; likelihood ratio, $-2\log\Lambda$ | 0.00 | 10.42 | 6.55 |
| Truncated geometric; likelihood ratio, $-2\log\Lambda$ | 0.00 | 2.32 | 1.08 |
| 1-way ANOVA intraclass correlation (Bartko, ref 8) | 1.0 | –0.23 | 0.34 |
| Spearman-Brown prediction correlation | 1.0 | –0.60 | 0.51 |
| 2-way ANOVA intraclass correlation (Bartko, ref 8) | 1.0 | 0.24 | 0.48 |
| Winer's anchor point correlation (Winer, ref 6) | 1.0 | 1.00 | 0.80 |

*All test statistics have 1 degree of freedom and are asymptomatically chi-square distributed.

It is trivial to show that all three data sets have a product-moment correlation of unity, whereas only set 1 reflects perfect agreement. Set 2 reflects an additive bias which is constant across items on the part of rater 2 relative to rater 1. Set 3, which reflects greater agreement between raters than set 2, displays a positive bias on the part of rater 2 which varies across items. Table 3 is a comparison of the models considered and of the various test statistics for the hypothesis of agreement between the two raters for the three data sets of table 2. For comparison purposes, this table includes four traditional intraclass correlations described by Bartko (7). For set 1, all of the models lead to the inference that the raters agree with each other. For sets 2 and 3 we note that the truncated Poisson model gives a likelihood ratio test statistic almost identical to that of the unrestricted Poisson model. Within the latter two families, the test statistics lead to rejection of the hypothesis of rater agreement for both sets, the descriptive significance levels each being $< 0.01$. As expected, the inference is much stronger for set 2 than for set 3. The truncated geometric model provides a less extreme picture of what is going on and serves to underline the relationship between model selection and subsequent inference. For set 2, the descriptive significance level is 0.13 and one could probably argue that for data of this type and size a nominal significance level much larger than the traditional 0.05 ought to be set. If this is so, we are led once again to reject the hypothesis of agreement.

For set 3, however, the descriptive significance level is 0.30, and we would infer that raters 1 and 2 are in agreement. Since the choice of model affects the inference so sharply here, one would be inclined to retrain the raters and repeat the experiment or augment the experiment with additional items to be rated. For example, we would not expect the inference to change if items were added whose scores were in no worse agreement, ie, a maximum spread of five units. However, if items which show greater disagreement are introduced, the force of the inference diminishes considerably. Thus, if an item scored 2 and 9 by raters 1 and 2 respectively were introduced, the likelihood ratio test statistic increases to 1.82, corresponding approximately to a significance level of 0.18. The addition of one more item, scored from 3 to 9, raises the test statistic to 2.08 with a corresponding approximate significance level of 0.15, and the hypothesis of agreement would very likely be rejected.

TABLE 4.—The No. of avoiding strategies observed by 9 raters independently judging the same 5 videotaped interviews (from an unpublished pilot study of the Psychosocial Oncology Group [RCS-01-76, Exploratory Study of Psychosocial Instrumentation])

| Interview (i) | Rater (j) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | $l_i$ |
| 1 | 1 | 1 | 3 | 0 | 2 | 0 | 1 | 2 | 0 | 10 |
| 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 3 | 1 | 16 |
| 3 | 1 | 1 | 3 | 4 | 3 | 1 | 1 | 4 | 1 | 19 |
| 4 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 15 |
| 5 | 0 | 2 | 0 | 1 | 1 | 2 | 1 | 2 | 2 | 11 |
| $c_j$ | 4 | 8 | 10 | 9 | 10 | 7 | 5 | 13 | 5 | 71 |

## Multiple Raters

Table 4 records the number of avoiding strategies observed by nine raters independently judging five videotaped interviews. These data are from a pilot protocol carried out by the Psychosocial Oncology Group. The number of avoiding strategies which is an integer from 0 to 9 inclusive is one dimension of the Coping Strategies Inventory, as modified and expanded by Penman (Psychosocial Oncology Group member) from Sidle et al (10).

For these data, the assumption of an unrestricted Poisson model seems reasonable in view of the fact that no more than four strategies were observed by any rater. The Pearson goodness-of-fit statistic given by equation [4] takes the value 8.73 with 8 degrees of freedom. For the companion statistic given by equation [3], we have $-2\log\Lambda = 8.84$, again on 8 degrees of freedom. Not unexpectedly, the two statistics are in excellent agreement, and imply that the nine raters are in agreement. On the other hand, if rater 8 had been in worse agreement, to the extent of a column total of $\geq 17$, we would have rejected the hypothesis of rater agreement at the 10% significance level.

## REFERENCES


1. LANDIS JR, and KOCH GG. The measurement of observer agreement for categorical data. Biometrics 33:159–197, 1977.
2. ———. An application of hierarchical Kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 33:363–374, 1977.
3. ———. A one-way components of variance model for categorical data. Biometrics 33:671–679, 1977.
4. FLEISS JL. Measuring nominal scale agreement among many raters. Psychol Bull 76:378–382, 1971.
5. LIGHT RL. Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull 76:365–377, 1971.
6. WINER BJ. Statistical Principles in Experimental Design, 2nd ed. New York, NY, McGraw-Hill, 1971.
7. BARTKO JJ. On various intraclass correlation reliability coefficients. Psychol Bull 83:762–765, 1976.
8. ———. The intraclass correlation coefficient as a measure of reliability. Psychol Rep 19:3–11, 1966.
9. BISHOP YM, FIENBERG SE, and HOLLAND PW. Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA, MIT Press, 1976.
10. SIDLE A, MOOS R, ADAMS J, ET AL. Development of a coping scale: a preliminary study. Arch Gen Psychiatry 20:(2)226–232, 1969.

# Clinical Trials in Psychosocial Medicine: Methodologic and Statistical Considerations. Part III. Assessing Measurement Techniques in Psychosocial Oncology [1]

Gary R. Morrow [2,3,*]

## SUMMARY

**The focus of the paper is on methodologic issues in assessing the validity of measurement techniques used in psychosocial oncology. Practical problems involved in assessing the validity of approaches, measures, and methods used in psychosocial oncology are outlined and briefly discussed. Techniques for partially addressing these problem areas are presented. Data from a study on the quantification of hope are examined in a multitrait-multimethod matrix assessment of construct validation which showed that little confidence can be placed in the validity of present methods of assessing "hope." Following a discussion of alternative statistical approaches to analyzing a multitrait-multimethod matrix for construct validity, the potential applicability of this approach to clinical oncology is presented through a hypothetical study on quantifying tumor types.**

[Cancer Treat Rep 64:451–456, 1980]

The physical and social sciences share the same principles of scientific investigation. Each, however, has its own set of methodologic problems surrounding the quantification of variables used in investigative studies. A detailed description of these differences from a theoretic framework is provided elsewhere (1). The purposes of this paper are: (*a*) to outline major practical problems encountered in investigating the validity of the methods, measures, and approaches used in psychosocial oncology; (*b*) to review specific approaches to some of these difficulties; and (*c*) to illustrate the application of a technique of construct validation, called multitrait-multimethod matrix analysis, that may prove as useful for clinical as it has for social scientists.

[1]Supported in part by Public Health Service grants CA-11198 and CA-19681 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare, and by the United Cancer Council, Inc, of Monroe County, NY.

[2]Psychosocial Medicine Unit, University of Rochester Cancer Center and Medical Center, NY.

[3]I gratefully acknowledge the assistance of Mike Feldstein (Division of Biostatistics and Epidemiology, Sidney Farber Cancer Institute, Boston, MA) for providing technical assistance, and Gail Drew, for typing the manuscript.

**Reprint requests to:* Dr. Gary R. Morrow, Department of Psychiatry, University of Rochester Medical Center, 300 Crittenden Blvd, Rochester, NY 14642.

## PRACTICAL PROBLEMS OF VALIDITY

Psychosocial oncology has moderately well-developed ancestry in clinical tools from allied fields such as psychology, psychiatry, behavioral medicine, social work, nursing, sociology, and other social sciences. Scientific investigation in this field is, however, in its relative infancy. As in all newly developed areas of investigation, the tools for measuring important variables are being borrowed and modified from allied disciplines or are being developed on their own.

No one has ever seen anxiety; no micro balance can measure social support; no fluorometric technique has been devised to quantify adjustment, and we have neither a TNM system for the classification of psychosocial adjustment nor a serum test for well-being. The search for an "emotional biopsy" of the cancer patient remains illusory. In short, the methods and techniques used for the assessment and quantification of constructs within psychosocial oncology require a technology different from that developed for the physical sciences.

The social sciences in general, and psychosocial oncology in particular, deal in large measure with hypothetical constructs. These constructs are sometimes one or more steps removed from physical or other measurable situations. A further problem is that sometimes constructs are used in idiosyncratic ways.

The construct of "adjustment" provides a brief example. Given two clinical cases, we probably could get reasonably good agreement on which patient demonstrated better "adjustment" to his treatment. However, were we to attempt to develop a measurement instrument to quantify what was meant by "adjustment" so that others could assess "adjustment" in a consistent manner, the task would prove more difficult.

In nuclear physics no one has ever directly observed a precursive strontium particle or neutron flux, yet these constructs justify their existence by behaving in a consistent and theoretically lawful manner. By showing such behavior they are measured with confidence in the validity of the measures. Although psychosocial oncology may be several years behind some of the clinical sciences in its available measures of demonstrated validity, progress is being made.

## GENERAL APPROACHES TO VALIDITY

Before focusing on particular techniques of instrument validation, it is useful to review applicable methods for the validation of the measurement tools presently being developed, with a particular emphasis on the usefulness of construct validation.

There is a fundamental difference between the social and physical sciences in the interpretation of validity. Physical scientists do not speak of validating a measurement procedure, technique, or instrument, but rather of validating scientific theory (2). A more constricted view is applied in the social sciences; validity involves assessing a measurement procedure in terms of the degree to which data arising from that specific procedure are consistent with the theoretic construct being investigated. Although there is some disagreement in nomenclature (3,4), there are three main approaches to the validation of a measuring procedure for an instrument used in the social sciences.

1. *Face validity.*—In the simple form, this, the most commonly used and also most commonly abused measure of validity, implies that an instrument "seems to or looks like" it measures a construct or a particular domain of interest. What is often misunderstood is that face validity is necessary but not sufficient for an assessment procedure. Face validity is essentially a judgment by consensus and is supported by no specific investigative, mathematic, or theoretic techniques.

2. *Criterion-related validity.*—Until the early 1950s, this measure—the matching of results from instruments or measuring procedures with an external observation or independent measure that served as a criterion—was the standard within validity measurement. Prominent difficulties with this type of validity investigation center around the choosing of the external criteria (5).

3. *Construct validity.*—This measure is perhaps the most applicable to the validation of measuring procedures and instruments used in the social sciences (6). In theory, this type of validity requires that the operations of measurement (procedures, tests, instruments) be associated with the theoretic constructs they purport to measure. In practice, the demonstration of this association has been more complex.

An early and fundamental problem in the validation of hypothetical constructs and of measuring tests, inventories, and other assessment instruments has been the assessment of construct validity. Campbell and Fiske (7) proposed the first major advance in the development of assessment, pointing out that the assessment of construct validity involves a distinction between convergence and discrimination, or convergent and discriminant validity. Construct validation requires not only confirmation by independent measurement procedures (a common denominator of the other types of validation) but also a demonstration of uniqueness and separate identification.

In addition to showing that scores of a trait (construct) X as measured by method (or instrument) A are associated with scores of X as measured by a second separate and independent method or instrument B (convergence), it is necessary to show that X (as measured by A or B) is *not* associated with Y (a trait or construct it is intended to differ from) as measured by A or B (divergence).

## THE USE OF MULTITRAIT-MULTIMETHOD MATRIX ANALYSIS FOR CONSTRUCT VALIDATION

Campbell and Fiske proposed a technique for assessing construct validity. They showed that it is possible to examine both convergent and divergent validity within a matrix of the intercorrelations of two or more theoretically unrelated traits or constructs measured by two or more unrelated methods. This approach, which is called multitrait-multimethod (MT-MM) matrix analysis, has become one of the most widely used techniques of construct validation within the social sciences (8).

### Example of Multitrait-Multimethod Matrix

The lower diagonal half of the example (table 1) presents the intercorrelations of three traits or constructs (anxiety, hostility, and hope) measured by three independent methods (verbal report, observer-rated, and self-reported). The number in parentheses below each trait notation in the table heading refers to the individual test, instrument, or procedure within a given method used for measuring each trait (9–14); for example, Gottschalk and Gleser (9) describe the instrument used for assessing the trait of "anxiety" by the "verbal report" method.

These data are from a substudy of an ongoing investigation at Rochester concerning the psychosocial adjustment of patients with Hodgkin's disease as well as

the parents of such patients. The purpose of the substudy was to assess the degree to which the construct or trait "hope" was operative through the methods commonly used to assess it. Did it exist as a trait divergent from anxiety and hostility that converged across the three methods commonly used to measure it? Or, did the method of measurement determine the degree to which it was a separate trait? Was there hope for "hope"?

## Terms Used in Multitrait-Multimethod Analysis

Some basic nomenclature is helpful in understanding how the MT-MM matrix is analyzed by the Campbell-Fiske criteria (7,15). The three diagonal submatrices outlined by the right-angled triangles in table 1 are called *heterotrait-monomethod blocks*. These blocks present the correlations among the three traits evaluated by each of the three separate methods. Each element within a block is the correlation between different traits measured by the same method.

The three off-diagonal submatrices outlined by rectangles are called *heteromethod-heterotrait blocks*. These blocks present the correlations between different traits evaluated by different methods. The diagonals of these blocks are called *validity diagonals* since they are the correlations between the same trait measured by different methods. The validity diagonal values are enclosed by parentheses in table 1.

## Criteria for Construct Validity in Multitrait-Multimethod Analysis

Campbell and Fiske suggested four informal criteria for evaluating the degree to which data presented in a MT-MM matrix reflect convergent, discriminant, and hence, construct validity.

1. The correlation between two values of the same trait measured by different methods (validity diagonal value) should be both statistically significant and high enough to encourage further consideration. This is evidence of convergent validity.
2. The correlation between the *same trait* measured by *different methods* (validity diagonal value) should be higher than correlations between *different traits* measured by *different methods*. In other words, a validity value for a variable should be higher than the correlations obtained between that variable and any other having neither trait nor method in common (ie, values within its row and column within the *hetero*trait-*hetero*method block).
3. The correlation between two values of the *same trait* measured by *different methods* (validity diagonal value) should be greater than correlations between *different traits* measured by the *same method*. In other words, a variable should correlate more highly with an independent effort to measure the same trait than with measures designed to get at different traits which happen to use the same method (ie, values within its row and/or column within the *hetero*trait-*mono*method block).
4. A similar pattern of trait intercorrelations should be shown in both monomethod and heteromethod blocks.

The last three criteria apply to discriminant validity. Inspection of the matrix (table 2) reveals that, with only one of three significant validity diagonals (ie, $r = 0.28$ between verbally reported hope and observer-rated hope), evidence for convergent validity is less than over-

TABLE 1.—MT-MM matrix*

| | | Method of measurement | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Verbal report | | | Observer-rated | | | Self-reported | | |
| | | Anxiety (9)† | Hostility (9) | Hope (10) | Anxiety (11) | Hostility (11) | Hope (11) | Anxiety (12) | Hostility (13) | Hope (14) |
| Verbal report | Anxiety | 1.00 | | | | | | | | |
| | Hostility | 0.28 | 1.00 | | | | | | | |
| | Hope | 0.36 | 0.31 | 1.00 | | | | | | |
| Observer-rated | Anxiety | (0.25)‡ | 0.06 | −0.10 | 1.00 | | | | | |
| | Hostility | 0.24 | (0.08) | −0.16 | 0.47 | 1.00 | | | | |
| | Hope | 0.12 | 0.17 | (0.28) | 0.60 | 0.45 | 1.00 | | | |
| Self-reported | Anxiety | (−0.04) | 0 | −0.04 | (0.47) | 0.27 | 0.12 | 1.00 | | |
| | Hostility | −0.10 | (0.20) | 0.12 | 0.03 | (−0.15) | 0.15 | 0.05 | 1.00 | |
| | Hope | −0.15 | 0.10 | (0.10) | 0.25 | 0.08 | (−0.05) | 0.64 | 0.05 | 1.00 |

*Total No. of subjects = 35.

†Nos. in parentheses below each trait refer to ref No. for the individual test, instrument, or procedure within a given method used for measuring each trait (refs 9–14).

‡ Nos. in parentheses are validity diagonal values.

TABLE 2.—Analysis of MT-MM matrix by Campbell-Fiske criteria

| Campbell-Fiske criteria | Trait | Proportion matching* | Example† |
|---|---|---|---|
| Significant validity diagonal | Anxiety | 2/3 | |
| | Hostility | 1/3 | |
| | Hope | 1/3 | (0.28) ($P < 0.05$); (0.10) (NS); (–0.05) (NS) |
| Validity diagonal > heterotrait-heteromethod row *and* column | Anxiety | 9/12 | |
| | Hostility | 8/12 | |
| | Hope | 6/12 | (0.28) > –0.16, –0.10, 0.17, 0.12<br>0.12 > (0.10) > –0.15, –0.10, –0.04<br>0.25, 0.08, 0.15, 0.12 > (–0.05) |
| Validity diagonal > heterotrait-monomethod row *or* column | Anxiety | 1/6 | |
| | Hostility | 1/6 | |
| | Hope | 1/6 | (0.28) < 0.60, 0.45<br>0.05 < (0.10) < 0.64<br>0.05 = (0.05) < 0.64 |

*No. of correlations examined for a given criterion/No. of correlations meeting the criterion.
†NS = not significant.

powering. Hope further suffers because it fails to meet discriminant criteria 2 and 3 (seven of 18 met).

With hope for "hope" dwindling rapidly, it is useful to point out one of the perhaps already obvious difficulties with the original Campbell-Fiske criteria. They suffer from a lack of rigor and objectivity, and their interpretation tends to be subjective. The primary analysis is the ratio of the number of correlations meeting a given criterion to the number of correlations examined for that criterion. Furthermore, it is not clear whether all four suggested criteria must be met.

## Alternative Approaches of Multitrait-Multimethod Matrix Analysis

At least five techniques have been proposed for an examination more statistically rigorous than that which the Campbell-Fiske criteria provide of the degree to which data in an MT-MM matrix support the construct validity of trait(s) examined (16–18). Table 3 compares these approaches with the Campbell-Fiske method. An excellent review of the techniques applied to a common numerical example has been written by Schmitt et al (15).

One major advantage of the path and factor analytic approaches presented in table 3 is that separation of method and trait variance is possible. This separation provides an estimate of the degree to which values of a given trait are dependent on (or independent of) the method of measurement used.

## An Application of Clinical Oncology

It is possible to construct an MT-MM matrix of raters and observations in place of methods and traits. This might make the technique useful, for example, to the clinical oncologist who wishes to assess the validity of a new system for quantifying tumor types by some schema of independent observations. Table 4 shows the structure of an MT-MM matrix obtained from a hypothetical study where three clinical raters (method) rated a number of tumors on the characteristics of size, angularity, and consistency (traits or constructs).

It is perhaps important to note that, theoretically, the number of methods or raters and the number of traits or constructs may be any number > 2. The practical issue, the sorting out of the statistical values, can become burdened with path analysis, since the estimated path coefficients are decomposed elements of the original correlations which are recombined to check the adequacy of the solution. A four-method by seven-trait MT-MM matrix will yield, for example, 84 estimated path coefficients.

As with the example described previously, the values in each row and column would be the respective intercorrelations. One of the analytic methods shown in table 3 would make it possible to determine: (*a*) the degree to which the characteristics were truly independent variables, or whether a lesser number of variables would be as useful, and (*b*) the degree to which the raters were consistent in their use of the characteristics.

Results of a path analysis could show if various raters agreed on how they used the characteristics in rating, and could address such questions as, "Do all raters agree among themselves or are certain raters rating idiosyncratically?" and "Are some characteristics highly interrelated?"

The practical problems involved in validating measures for use in psychosocial oncologic clinical trials may seem parochial, but the more we wrestle with them, the more potentially applicable approaches we find for their solution.

One important way to increase the probability of finding solutions is by talking to researchers from other scientific fields. The approaches to validity presented here

TABLE 3.—Summary of analytic methods for MT-MM matrix and the degree to which various objectives are achieved*

| | Method of analysis | | | | | |
|---|---|---|---|---|---|---|
| | Inspection | ANOVA | Path-analytic | Factor analytic | | |
| Purpose | Campbell and Fiske (ref 7)† | Kavanagh et al (ref 19) | Werts and Linn (ref 20) | Golding and Seidman (ref 16) | Jackson (ref 17) | Tucker (ref 18) |
| Unconfounding trait and method variance | No | No | Various hypothesized models of sources of variance can be tested when N is reasonably large | Common and specific variance are confounded in first transformation | Common and specific variance are confounded in first transformation | Solution rests on satisfactoriness of communality estimates and decision on number of factors to rotate |
| Possibility of evaluation of criteria for individual traits and methods as opposed to factors or components | Yes | No, there is a single summary figure for traits and methods | Yes, each measured trait is hypothesized to be a result of trait, method, and unique variance | No, convergence of trait factors across methods is only objective | Convergence and discrimination are at level of factor scores, though rotations are targeted for individual measured variables | Estimation of intercorrelation of subject, method, and trait factors is the objective |
| Basis for evaluating degree of convergent and discriminant validity | % of time greater than/less than relationships are correct | Tests of significance for trait and method convergence | Chi-square test for overall fit of hypothesized models | Emergence of clear trait factors in second-step analysis | Emergence of trait and method factors in second-step analyses | Levels of intercorrelation among subject, trait, and method factors in core matrix |
| Location of individuals on an underlying trait | No | No | Yes | Yes | Yes | Yes |

*Modified by permission of Schmitt et al (ref 15) and *Multivariate Behavioral Research*.
†Source of example.

TABLE 4.—Hypothetical multitrait-multirater matrix

| | Rater 1 | | | Rater 2 | | | Rater 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Size | Angularity | Consistency | Size | Angularity | Consistency | Size | Angularity | Consistency |
| Rater 1 | | | | | | | | | |
| Size | 1.00 | | | | | | | | |
| Angularity | $r_{S1A1}$* | 1.00 | | | | | | | |
| Consistency | $r_{S1C1}$ | $r_{A1C1}$ | 1.00 | | | | | | |
| Rater 2 | | | | | | | | | |
| Size | ($r_{S1S2}$) | $r_{A1S2}$ | $r_{C1S2}$ | 1.00 | | | | | |
| Angularity | $r_{S1A2}$ | ($r_{A1A2}$) | $r_{C1A2}$ | $r_{S2A2}$ | 1.00 | | | | |
| Consistency | $r_{S1C2}$ | $r_{A1C2}$ | ($r_{C1C2}$) | $r_{S2C2}$ | $r_{A2C2}$ | 1.00 | | | |
| Rater 3 | | | | | | | | | |
| Size | ($r_{S1S3}$) | $r_{A1S3}$ | $r_{C1S3}$ | ($r_{S2S3}$) | $r_{A2S3}$ | $r_{C2S3}$ | 1.00 | | |
| Angularity | $r_{S1A3}$ | ($r_{A1A3}$) | $r_{C1A3}$ | $r_{S2A3}$ | ($r_{A2A3}$) | $r_{C2A3}$ | $r_{S3A3}$ | 1.00 | |
| Consistency | $r_{S1C3}$ | $r_{A1C3}$ | ($r_{C1C3}$) | $r_{S2C3}$ | $r_{A2C3}$ | ($r_{C2C3}$) | $r_{S3C3}$ | $r_{A3C3}$ | 1.00 |

*Example: $r_{S1A2}$ correlation between value of size given by Rater 1 and value of angularity given by Rater 2.

are not the sole property of a single discipline. Often their conception, birth, and development have been combined efforts of multiple parents. For example, the MT-MM analysis to construct validity was conceived within psychology, but it was not until the several alternative statistical methods of analysis were shown to be applicable that this method began to appear in the literature of such disciplines as psychology, sociology, child development, and political science.

The multidisciplinary sharing of ideas and techniques has proved useful in many scientific areas, and conferences such as this, where that sharing of ideas can take place, are of extraordinary importance.

## REFERENCES


1. TOGERSON WS. Theory and Methods of Scaling. New York, NY, John Wiley, 1958.
2. FRANK P, ed. The Validation of Scientific Theories. Boston, MA, Beacon, 1956.
3. CRONBACK LJ. Test validation. *In* Educational Measurement, 2nd ed (Thorndike RL, ed). Washington, DC, American Council on Education, 1971, pp 443–508.
4. AMERICAN PSYCHOLOGICAL ASSOCIATION. Standards for Educational and Psychological Tests. Washington, DC, 1966.
5. MORROW GR. The relationship between psychological and physiological measures of anxiety. Psychol Med 8:95–101, 1978.
6. MESSICK S. The standard problem: meaning and values in measurement and evaluation. Am Psychol 10:955–966, 1975.
7. CAMPBELL OT, and FISKE DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 56:81–105, 1959.
8. NEALE JM, and LIEBERT RM. Science and Behavior: An Introduction to Methods of Research. New York, NY, Prentice-Hall, 1973, pp 119–122.
9. GOTTSCHALK LA, and GLESER GC. The Measurement of Psychological States Through the Content Analysis of Verbal Behavior. Berkeley, CA, University of California Press, 1969.
10. GOTTSCHALK LA. A hope scale applicable to verbal samples. Arch Gen Psychiatry 30:779–785, 1974.
11. DEROGATIS LR. The SCL-90 Manual. Baltimore, MD, Johns Hopkins Press, 1977.
12. SPIELBERGER CD, GORSUCH RL, and LUSHENE R. The State-Trait Anxiety Inventory (STAI). Palo Alto, CA, Consulting Psychologists Press, 1968.
13. KAUFMANN H. Correlations among verbal response hostility measures. Percept Mot Skills 20:258, 1965.
14. BECK AT, WEISSMAN A, LESTER O, ET AL. The measurement of pessimism: the hopelessness scale. J Clin Consult Psychol 42:861–865, 1974.
15. SCHMITT N, COYLE BW, and SAARI BB. A review and critique of analyses of multitrait-multimethod matrices. Multivariate Behav Res 12:447–478, 1977.
16. GOLDING SL, and SEIDMAN E. Analysis of multitrait-multimethod matrices: a two step principal components procedure. Multivariate Behav Res 9:479–496, 1974.
17. JACKSON ON. Multimethod factor analysis in the evaluation of convergent and discriminant validity. Psychol Bull 29:259–271, 1969.
18. TUCKER LR. Some mathematical notes on three-mode factor analysis. Psychometrica 31:279–311, 1966.
19. KAVANAGH MJ, MACKINNEY AC, and WOLINS L. Issues in managerial performance: multitrait-multimethod analyses of ratings. Psychol Bull 75:34–39, 1971.
20. WERTS CE, and LINN RL. Path analysis: psychological examples. Psychol Bull 74:193–212, 1970.

# Evaluation of Toxicity: Clinical Issues [1]

**Teresa J. Vietti** [2,*]

## SUMMARY

**Toxicity criteria used by three of the major cooperative groups are reviewed and discussed. The definition of acceptable toxicity will depend on the extent of disease and the availability of effective therapy. For some forms of toxicity there are adequate early warning signals; for other forms no satisfactory criteria yet exist. Until recently, no attempt had been made to systematically evaluate the quality of the patient's life during and after therapy. We are just now beginning to assess the long-term pathophysiologic changes that occur during and after multimodal therapy. The protocol must include information concerning agent toxicity, criteria for modification of therapy, and time intervals for documentation of toxicity. We must continue to develop therapeutic regimens which control the disease but interfere minimally with the quality of life.**

[Cancer Treat Rep 64:457–461, 1980]

During the past 10 years the evaluation of toxicity has undergone significant revisions by the cooperative groups. Standard criteria have been developed based on symptomatology, physical findings, and laboratory data. Toxicity is assessed and monitored (*a*) to determine whether the patient is receiving adequate therapy, (*b*) to evaluate the toxicity associated with a given therapeutic regimen, (*c*) to compare toxic effects of various regimens, and (*d*) to predict unacceptable, intolerable, or life-threatening toxicity.

## TOXICITY CRITERIA

### Acute Pathophysiologic Changes

In the following discussion, I will comment on the toxicity criteria which are in use by three of the major cooperative groups.[3] Since their criteria for grading toxicity have not been published, I will rank toxicity from grades 1 to 4, designating increasing degrees of pathophysiologic abnormalities according to the following parameters:

*grade 0.*—No detectable change.

*grade 1.*—Minimal changes that do not influence the patient's general activity. Laboratory data are not significantly decreased below normal values and do not usually require any corrective therapy, but do show evidence of biologic activity by the agent.

*grade 2.*—Moderate changes which interfere with the patient's activity. Laboratory data are significantly lower than normal. Minor supportive measures may be required but therapy is tolerable.

*grade 3.*—Severe changes which are significant and require corrective measures, or a delay or interruption of therapy.

*grade 4.*—Life-threatening, unacceptable, and intolerable toxicity. Therapy, if resumed, must be decreased so that this degree of toxicity will not recur. This form of toxicity may also be irreversible and/or fatal.

These ratings of toxicity are merely guidelines, and their use will depend on the stage of the patient's disease and predicted response to the therapeutic regimen. Thus, grade 3 toxicity may be acceptable only in a patient whose prognosis is poor and/or when it is known that response will not be optimal otherwise. In the treatment of a child with stage I Wilms' tumor, an absolute neutrophil count of $< 1000/\mu l$ is not acceptable, whereas it might be ac-

[1]Supported in part by Public Health Service grant CA-05587 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare, and by the Fern Waldman Memorial Fund.

[2]Edward Mallinckrodt Department of Pediatrics, Washington University School of Medicine, and the Division of Hematology and Oncology, St. Louis Children's Hospital, MO.

[3]Toxicity criteria of the Southwest Oncology Group, Cancer and Leukemia Group B, and the Eastern Cooperative Oncology Group. Personal communication.

**Reprint requests to*: Dr. Teresa J. Vietti, Edward Mallinckrodt Department of Pediatrics, Washington University School of Medicine, 500 S Kingshighway, PO Box 14871, St. Louis, MO 63178.

TABLE 1.—Comparison of toxicity criteria

| Toxicity | Group | Grade— 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Wbc count (cells × $10^3$/μl) | X | 3–<4 | 2–<3 | 1–<2 | <1 |
| | Y | 3–<4.5 | 2–<3 | 1–<2 | <1 |
| | Z | 3–<4 | 2–<3 | 1–<2 | <1 |
| Granulocyte count (cells × $10^3$/μl) | X | 1.5–<2.0 | 1.0–<1.5 | 0.5–<1.0 | <0.5 |
| | Y | 1.5–<1.9 | 1.0–<1.5 | 0.5–<1.0 | <0.5 |
| | Z | 1.0–<1.5 | 0.5–<1.0 | 0.25–<0.5 | <0.25 |
| Platelet count (cells × $10^3$/μl) | X | 100–<150 | 75–<100 | 25–<75 | <25 |
| | Y | 90–<130 | 50–<90 | 25–<50 | <25 |
| | Z | 75–<100 | 50–<75 | 25–<50 | <25 |
| Infection | X | Mild, no fever | Fever, <38°C | Sepsis, >38°C | Sepsis, shock |
| | Y | <38°C | >38°C | Chills, >40°C | Fever, shock |
| | Z | — | — | — | — |
| Nausea and vomiting | X | Nausea only | Both, occasionally | Both <2 times/day | Severe, uncontrolled |
| | Y | Nausea only | Both, occasionally | Both, incontrolled | — |
| | Z | Nausea only | Both, <6 times/day | Both, uncontrolled | — |
| Diarrhea | X | Semi-solid | 3–4 liquid | >4 liquid | Hemorrhagic, dehydration |
| | Y | No dehydration | Dehydration | Grossly bloody | — |
| | Z | 3–4, no dehydration | >4, needs iv | Bloody, needs iv or blood transfusion | — |
| BUN level (mg/dl) | X | 21–30 | 31–50 | >50 | Uremic coma |
| | Y | 21–40 | 41–60 | >60 | Clinical uremia |
| | Z | 21–40 | 41–60 | >60 | — |
| Bilirubin level (mg/dl) | X | — | — | — | — |
| | Y | 1.5–2 × normal | 2.5–5 × normal | >5 normal | Coma |
| | Z | 1–2 | 2.1–5 | >5 | — |

ceptable for a child with stage IV Wilms' tumor with an unfavorable (anaplastic) histology.

Table 1 compares the toxicity grading criteria of these three groups. The toxicity criteria for total wbc count are similar for the three groups. Group Z has a much lower total absolute neutrophil count value for intolerable (grade 4) toxicity.

Toxicity may be compounded and thus greatly increase the possibility of complications developing during therapy. If an agent causes mucosal ulceration as well as myelosuppression, the likelihood of an infection developing is greatly increased.

Another problem area involves patients who have received prior extensive myelosuppressive therapy and have compromised bone marrow reserves. These patients may have normal blood cell counts but may have severe granulocytopenia while receiving subsequent courses of chemotherapy which are well-tolerated by patients with normal bone marrow reserves.

The ratings of myelotoxicity reflect only the nadir of myelotoxicity and not the duration. Thus, a transient (< 1 week) decrease in the absolute neutrophil count to < 500/μl may be acceptable toxicity, but if it persists for a longer period of time it is not acceptable because of the likelihood that the patient will develop a fulminating infection. In acute leukemia it is difficult to evaluate the myelotoxicity during remission induction because marrow function is already severely compromised. Masera and Carnelli (1) have shown that the total wbc count nadir for combination vincristine and prednisone is the same as for combination vincristine, prednisone, and daunorubicin, but the duration of myelotoxicity is significantly longer (fig 1). Only groups X and Y have a grading criteria for infection. It may be necessary to reconsider the use of infectious complications in the grading of toxicity.

A brief comment on the toxicity criteria for thrombocytopenia is necessary. A normal platelet count should be between 140 and 440 × $10^3$/μl. The values used by group Z to indicate grade 1 toxicity are too low.

The toxicity criteria for nausea and vomiting and diar-

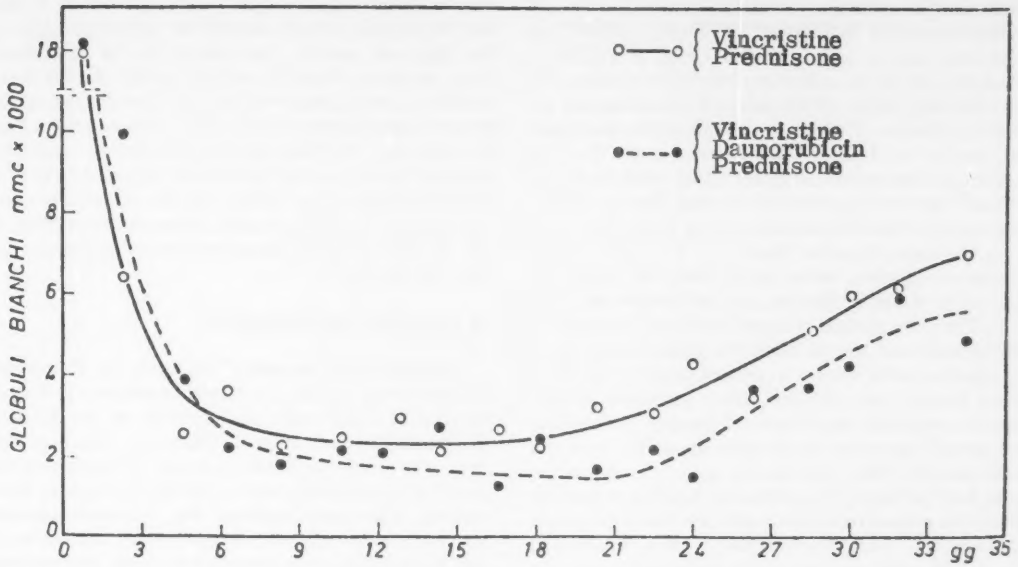



FIGURE 1.—Comparison of the nadir and duration of leukopenia in children with acute lymphocytic leukemia receiving 2 different remission-induction regimens. Vertical axis = total wbc count × $10^3/\mu l$ and horizontal axis = No. of days after initiating therapy. Although the nadir of the 2 regimens is the same, the duration of leukopenia is significantly longer in the children receiving vincristine, prednisone, and daunorubicin (reproduced *from* ref 1).

rhea that are used by the three cooperative groups are also shown in table 1. Note that nausea and vomiting not controlled by medication is designated grade 3 by groups Y and Z. Admittedly, the clinician does not consider this toxic effect serious, but as it becomes rapidly unacceptable to the patient, therapy must be stopped or modified. We should consider interference with the patient's lifestyle and quality of life when we grade toxicity.

Only group X rated bloody diarrhea as grade 4 or life-threatening. Since the development of bloody diarrhea probably means mucosal ulceration that could lead to perforation and peritonitis, the ranking of grade 4 for this toxic effect is probably justified. In addition, not many patients will tolerate corrective measures (hydration and blood transfusions) on a long-term basis. Incidentally, any change of normal bowel habits due to toxicity should probably be rated as grade 1.

Table 1 also gives the toxicity ranking for BUN and bilirubin by the three groups. A BUN level of 50 mg/dl is an ominous indication of serious toxicity and requires immediate cessation of therapy, with its resumption at a lower dose or other preventive measures once the BUN returns to normal values. Thus, according to the criteria which I have suggested, this should be rated as grade 4 toxicity. The same statements regarding toxicity can also be said for the bilirubin levels.

Certain forms of toxicity are regarded as early warning signs calling for cessation, reduction and/or delay in therapy, and/or institution of preventive therapy. The development of dysuria with microscopic hematuria secondary to cyclophosphamide therapy is considered grade 1 toxicity, but requires cessation of therapy until signs of cystitis disappear. When cyclophosphamide therapy is resumed, additional measures are taken to ensure rapid excretion of the active cyclophosphamide metabolite.

There are some forms of toxicity which may be life-threatening or even fatal but for which we have no predictive early warning. Almost all anticancer agents, including radiation therapy, produce immunosuppression which may lead to severe, overwhelming infection. No satisfactory warning signs have been developed for myocardiopathy due to the anthracyclines or the increased cardiovascular disease in patients with prostatic cancer receiving estrogen therapy. Byar (2) showed that the increased incidence of cardiovascular mortality secondary to estrogen therapy offsets the decrease in mortality due to prostatic cancer in patients entered in a Veterans' Administration Cooperative Group study.

## Quality of Life

Psychosocial problems that may develop as a result of our therapy are almost completely unexplored areas. For example, some forms of toxicity do not disturb the phy-

sician, but the patient may consider them intolerable and refuse to continue his therapy. Dr. Jimmy Holland[4] has developed Cancer and Leukemia Group B guidelines which provide for the collection of serial observations, by the staff and patient, of the patient's behavioral and attitudinal changes. Patients are monitored for psychosocial function by three methods of assessment. The first method includes an overall global rating which ranks the patient's level of psychosocial function despite illness. The second method is a handicap rating scale which evaluates the level of patient function in five spheres: employment limitation, restriction of social life, emotional instability, cognitive function, and self awareness of illness. The third method is a profile of mood states by a self-administered test in which the patient rates his or her emotions with respect to tension, anxiety, hostility, vigor, fatigue, and confusion. These guidelines are designed to establish which forms of therapy are tolerable and permit the patient to maintain acceptable work and social activity. They can also be used to compare the mean level of impaired psychosocial function of patients entered in various treatment regimens. These guidelines would be very useful, for example, in the assessment of the toxicity of maytansine, a new anticancer agent that is associated with lethargy, generalized malaise, and depression. In clinical trials with maytansine, a significant number of patients have refused further therapy.

### Long-Term Complications

The toxicity criteria that have been developed were designed to assess and monitor the acute toxicity induced by chemotherapeutic agents, but they seem to be equally applicable to the acute toxicity that develops with radiation therapy used alone or combined with chemotherapy. We are just now beginning to address ourselves to the development of some standard criteria for the assessment of long-term complications of therapy. These criteria will necessarily include complications such as organic brain damage, flexion contracture of the extremity, interference with normal bowel function, secondary tumors, and the long-term psychosocial sequelae that develop as a result of our therapy.

## PROTOCOL DESIGN: ASSESSMENT AND MONITORING OF TOXICITY

The background section of each protocol should include information concerning the expected toxicity of the therapeutic regimen with respect to type, degree, and duration. The section which discusses the agents to be used should include the pharmacology of the drug with respect to metabolic activation, inactivation, and excretion. It should enumerate the drugs that will enhance toxicity or may interfere with the agent's antitumor effect. If age, sex, or pathophysiologic status of the patient will enhance the observed toxicity, this should also be mentioned. Also, provision should be made to modify therapy if excessive toxicity is observed. Because of erratic absorption of some agents administered orally, increased doses may be indicated. The time intervals selected for documentation of toxicity should include the expected nadir. An interim analysis of the toxicity and the response to therapy should be performed so that therapeutic modification can be made to avoid excessive toxicity or perhaps inadequate therapy.

## IS TOXICITY NECESSARY?

Is toxicity really necessary? Obviously it is, if moderate or even severe toxicity is required to achieve a prolonged survival or a high cure rate. A study at the St. Jude Children's Research Hospital (Memphis, TN) (3) of children with acute lymphatic leukemia indicated that the duration of remission was significantly longer in those patients who were receiving the maximum-tolerated amount of combination 6-mercaptopurine, methotrexate, and cyclophosphamide compared to those who received only one half of the maximum-tolerated doses. Subsequent studies (4) showed that the addition of cyclophosphamide or cyclophosphamide plus cytosine arabinoside to 6-mercaptopurine plus methotrexate greatly increased the incidence of complications but did not increase the number of long-term survivors. The Pediatric Division of the Southwest Oncology Group is evaluating the degree of myelosuppression necessary to sustain prolonged remissions and high cure rates in children with acute lymphatic leukemia. In this study, the wbc count will be maintained by dose modification at a level of between 3000 and 4500/μl in one group and of between 3000 and 4500/μl in another group. In the treatment of stage I Wilms' tumor (4) it has been demonstrated that radiation therapy is not necessary if the tumor has been completely excised.

## CONCLUSION

In all of our studies, the first objective must be to control the patient's disease. As we achieve this objective, we must then determine how to modify therapy in order to minimize toxic complications and to improve the quality of life of the patient during and after therapy.

## REFERENCES


1. Masera G, and Carnelli V. Terapia della leucemia acuta dell'infanzia con la nuova associazione daunomicina-vincristina-prednisone. Minerva Pediatr 21:39–46, 1969.
2. Byar DP. The Veterans Administration Cooperative Urological Research Group's studies of cancer of the prostate. Cancer 32:1126–1130, 1973.


[4]Holland J. Personal communication.

3. SIMONE JV. Factors that influence haematological remission duration in acute lymphocytic leukaemia. Br J Haematol 32:465–472, 1976.

4. D'ANGIO GJ, EVANS AE, BRESLOW N, ET AL. The treatment of Wilms' tumor. Results of the National Wilms' Tumor Study. Cancer 38:633–646, 1976.

# Evaluation of Toxicity: Statistical Considerations [1]

**Jay Herson** [2,*]

## SUMMARY

**This paper presents some statistical considerations in the evaluation of toxicity in phase III clinical trials in cooperative oncology groups. Analysis is concentrated on the frequency of unacceptable toxicity by treatment period within homogeneous prognostic groups. Relevant toxicity questions are found to be of two types: general and specific. The general questions are common to all trials and either are descriptive or relate to the assessment and etiology of toxicity. The specific questions are built into the design of the study and some are unique to phase III trials. Various aspects of design and analysis are discussed for each type of question, including application of life-table methods for analysis, guidelines for the creation of toxicity indices, and an evaluation of various approaches to answering the fundamental toxicity question in cancer therapy, "Is toxicity really necessary?" The role of toxicity data in final treatment evaluation and the methods for the assessment of late effects of therapy are discussed briefly.**

[Cancer Treat Rep 64:463–468, 1980]

The role of toxicity evaluation in phase I and II clinical trials in cancer appears well-understood. Only a superficial reference to toxicity is found in the design and analysis of phase III trials, where response is considered the important endpoint. There are two reasons why statistical analysis of toxicity data is necessary in phase III trials. First, these trials frequently present combination chemotherapy and/or multimodal treatments for which toxicity has never been evaluated. Second, toxicity data can suggest improvements or possible alternatives in therapy to make survival more comfortable for the patient, while making only the slightest compromise with response.

This paper presents some statistical considerations in the evaluation of toxicity in phase III clinical trials in cooperative oncology groups. The first section covers data collection and toxicity description. The section on design and analysis covers the two categories of toxicity questions that are asked in a clinical trial: the general types which apply to all trials and the specific types which are built into the design of the trial. Several aspects of design and analysis will be discussed for each type. The role of toxicity data in final treatment evaluation and the description of late effects of therapy will also be discussed.

[1]Supported in part by Public Health Service grant CA-12014 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]Department of Biomathematics, M. D. Anderson Hospital and Tumor Institute, The University of Texas System Cancer Center, Houston.

**Reprint requests to:* Jay Herson, PhD, Department of Biomathematics, M. D. Anderson Hospital and Tumor Institute, The University of Texas System Cancer Center, Houston, TX 77030.

## DATA COLLECTION AND TOXICITY DESCRIPTION

### Preliminary Planning

At the outset of a clinical trial, decisions must be made as to what toxicity data are to be collected, how toxicity will be graded, and how the data will be collected, stored, processed, and analyzed. Investigators must define a *minimum* required toxicity data set, assess the ability of all institutions to collect the data at required precision, and describe the use of each item of the data set in interim and final data analyses. If a research team cannot describe the analytic use of an item *at the commencement* of a trial, they should not require its collection *during* the trial.

Not all data need be required of every patient at every institution. For example, if cellular and humoral measures of immune response are to be developed, only those institutions with the facilities for this kind of work need

participate until the value and practicality of collecting these data are determined.

### Analytic Definitions

In order to evaluate toxicity, three analytic definitions must be made at the beginning of a study: treatment periods, toxicity groups, and prognostic groups for response.

*Treatment periods.*—The trial should be divided into time periods which correspond to important treatment events, eg, a course of chemotherapy, reinforcement therapy, several courses of chemotherapy including one course of radiotherapy, etc. One of the fundamental challenges will be to summarize toxicity experience during each treatment period. The periods should not be so broad that interpretation will be difficult, nor so narrow that data storage becomes prohibitive and analysis is highly vulnerable to missing data.

*Toxicity groups.*—Definitions should be drawn up for three toxicity groups: those toxic effects that are *biologic effects* (grades 1–3), those that are *early warnings* of further serious sequelae (grades 1–3), and those that are *life-threatening or unacceptable* (grade 4).

*Prognostic groups for response.*—Previous studies of many diseases have shown that patients may be grouped by data collected at the start of treatment into groups homogeneous in response prognosis. Where possible, these prognostic groups should be defined.

## DESIGN AND ANALYSIS FOR TOXICITY QUESTIONS

Two types of questions are asked about toxicity in oncology clinical trials: *general* and *specific*. While the general questions are common to all trials, the specific questions are inherent in the design of the trial.

### General Questions

The two types of general questions are those that are *descriptive* of toxicity and those that relate to the *assessment and etiology* of toxicity. Those questions that are *descriptive* of toxicity (to be addressed in interim and final data analyses) are:

*a.* Which patients experience toxicity?
*b.* What types of toxicity are experienced?
*c.* What is the severity of the toxicity?
*d.* When do the toxic effects occur during therapy?

These questions may be addressed in a table (table 1) that is both treatment-group and prognostic-group specific. The hazard (incidence) and cumulative incidence of first episode biologic effect and unacceptable toxicity are presented by treatment period. These estimates are computed by life-table methods (1). A percentage distribution of each toxicity group by specific toxicity type is included. For comparison purposes this table includes a response summary (in this case remission duration and survival time). The cumulative unacceptable toxicity incidence

TABLE 1.—Toxicity description for treatment 1 (good prognosis patients)

| | Treatment period | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| No. of— | | | | | |
| Cumulative wks | 4 | 7 | 10 | 13 | 16 |
| Evaluable patients | 80 | 80 | 77 | 74 | 70 |
| Treatment modifications | 3 | 4 | 4 | 2 | 6 |
| Toxicity (grade 1–3) | | | | | |
| Hazard/100 | 16.0 | 19.2 | 12.4 | 11.4 | 15.4 |
| Cumulative incidence/100 | 15.0 | 29.1 | 38.1 | 44.8 | 52.7 |
| Toxicity types (%)—All | 100 | 100 | 100 | 100 | 100 |
| Leukopenia | 74.6 | 76.1 | 72.6 | 72.0 | 81.7 |
| Thrombocytopenia | 7.1 | 5.8 | 5.1 | 4.7 | 6.2 |
| Genitourinary | 2.0 | — | — | 1.7 | 0.5 |
| Hepatic | — | — | 0.8 | — | — |
| Other | 16.3 | 18.1 | 22.3 | 21.6 | 11.6 |
| Toxicity (grade 4) | | | | | |
| Hazard/100 | 1.3 | — | 2.0 | — | 2.0 |
| Cumulative incidence/100 | 1.3 | 1.3 | 3.3 | 3.3 | 5.3 |
| Toxicity types (%)—All | 100 | 100 | 100 | 100 | 100 |
| Leukopenia | 100 | — | 50 | — | 100 |
| Other | — | — | 50 | — | — |
| Response summary | | | | | |
| % of patients in remission | 100 | 92.1 | 87.5 | 83.1 | 78.2 |
| % of patients surviving | 100 | 100 | 99.2 | 99.2 | 98.1 |

estimates can be used to indicate the appropriateness of early termination of a treatment.

A major challenge for clinicians and statisticians is the development of measures of summarizing toxicity experience during each treatment period.

Leukopenia is generally defined by some function of the wbc count. This function might be called a *toxicity index*. Figure 1 illustrates what the repeated measures of wbc in one patient might look like during a treatment period. A common index of leukopenia is "wbc count nadir of $\leq 1500/mm^3$ during the period." Although this is not a textbook order statistics problem, it is expected that the distribution of the nadir would depend on the number of repeated measures taken during the period. This number may vary from patient to patient and may be related to toxicity experience. Breslow and Zandstra (2) encountered a similar record-length bias in applying a definition of lymphocytosis based on a maximum lymphocyte count observed during remission. Elsewhere in this issue, Dr. Vietti discusses the clinical significance of the duration of a "low" wbc count (eg, $< 2000/mm^3$). This index would probably not be as vulnerable to record-length bias as the nadir. An important consideration is what number and what frequency of repeated wbc measures are needed to estimate the duration of the nadir. Estimation of duration might be enhanced by incorporating clinical data on infection into the estimation rule. Other candidates for a toxicity index might be the median wbc to nadir wbc ratio and the area under the wbc curve.

Several general considerations in the creation of indices are:

1. an index should be relatively uncorrelated with other toxicity indices,
2. indices should be robust to occasional missing data and occasional extreme values,
3. the user should determine the contribution of measurement error to the variance of the indices, and
4. the user should know the effect of regression toward the mean on inference.

Description might be facilitated by the creation of toxicity indices which are functions of several measures such as neutrophils, lymphocytes, polymorphonuclear cells, blast cells, platelets, hemoglobin, infection, etc. These

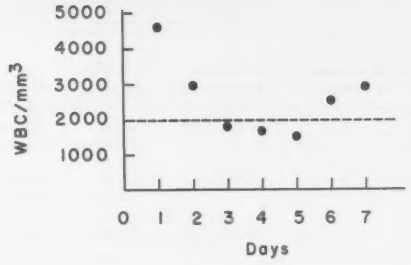



FIGURE 1.—Repeated measures of wbcs/$mm^3$ in 1 patient during a treatment period.

indices might reflect clinical processes such as immunosuppression and myelosuppression. The method of principal components (3) might be a useful starting point for the creation of such indices. Table 2 presents an application of principal components analysis using data from leukemic children in remission. The second index, $L_2$, which explains 26.8% of the variance, may be considered an index of bone marrow function.

Those questions that relate to the *assessment and etiology* of toxicity (to be addressed in interim and final data analysis) are:

*a.* What was the duration of toxicity?
*b.* Are there repeated episodes of the same toxicity? Are the effects cumulative?
*c.* Does the experience and/or management of the toxicity affect response?
*d.* Can unacceptable toxic effects be anticipated from "risk factors" (patient characteristics and early signs and symptoms)?

These questions may be addressed by reporting on each patient with unacceptable toxicity at each interim data analysis following the initial episode of toxicity. Table 3 shows the types of data to be reported. A similar table was used by Vaitkevicius et al (4) in the study of 5-fluorouracil.

At the end of a trial the effect of various risk factors on toxicity can be examined by comparing patients with unacceptable toxicity to the rest of their prognostic group

TABLE 2.—Principal components

| Toxicity index | Coefficients | | | | | | % variance |
|---|---|---|---|---|---|---|---|
| | Wbcs | Lymphocytes | Polymorphonuclear cells | Blast cells | Platelets | Hemoglobin | |
| $L_1$ | 0.597 | 0.278 | 0.329 | 0.552 | 0.037 | 0.390 | 37.6 |
| $L_2$ | –0.326 | 0.408 | 0.432 | –0.418 | 0.460 | 0.395 | 26.8 |
| $L_3$ | 0.117 | –0.457 | –0.091 | 0.187 | 0.848 | –0.122 | 14.1 |
| Total | | | | | | | 78.5 |

TABLE 3.—Toxicity assessment and etiology of patients with grade 4 toxicity

| Patient No. | Treatment arm | Institution location | Prognosis group | Response | Type of toxicity | Duration | Early signs | Management supportive care | Management dose |
|---|---|---|---|---|---|---|---|---|---|
| 58016 | 1 | Florida | Good | Remission (14 wks) | Leukopenia | 6 days | None | — | Decreased |
| 57110 | 2 | Cleveland | Good | Remission (21 wks) | Infection | 11 days | Leukopenia | Hospital | Decreased |
| 54199 | 2 | Mississippi | Poor | Death (9 wks) | Infection | 3 days | CNS | Hospital | — |

in terms of prestudy patient characteristics and early signs and symptoms. Such an analysis might suggest a need for a retrospective study of previous trials to investigate further possible risks such as Down's syndrome as a factor in monilial and fungal infections in children being treated with chemotherapy for leukemia.

Large samples are needed if meaningful risk factor analyses are to be performed because unacceptable or grade 4 toxicity occurs infrequently and because the risk factors must be described with regard to specific toxic effects such as cardiomyopathy or life-threatening infection. Therefore, this type of analysis will be facilitated by combining data from several cooperative groups. The National Cancer Institute recently performed an intergroup analysis of risk factors in Adriamycin-induced congestive heart failure[3] in 4018 patients and daunomycin-induced cardiotoxicity in 5613 patients (5). For each study, the dosage-incidence curve is estimated and risk factors such as age, sex, radiotherapy, and previous heart disease history are investigated.

These studies point out the need for the development of standardized toxicity criteria, data collection, and data processing methods to facilitate analyses that pool data from several cooperative groups or institutions.

An important aim in assessment is the determination of how the experience and/or management of grade 4 toxicity might affect response. "Management" refers not only treatment interruption and dosage change, but also the use of drugs such as lithium and mannitol to treat toxicity. This question addresses the effect of a patient entering a transient state (toxicity) on response time. Several investigators have approached this problem while dealing with the effect of heart transplant on survival time. Turnbull et al (6) have compared both nonparametric and parametric approaches. Mantel and Byar (7) have applied Mantel-Haenszel-Cox life-table procedures to the problem. Crowley and Hu (8) would use Cox's life-table regression model to allow for the assessment of the effect of toxicity experience on response by modelling the change in relapse or survival hazard after the onset of toxicity. The model assesses the toxicity effect while adjusting for baseline characteristics such as treatment and prognostic group and other time-dependent covariates such as onset of lymphocytosis during remission.

## Specific Questions

The specific questions are inherent in the design of the clinical trial and many are unique to phase III clinical trials. These questions would be addressed in the final report and perhaps midway into the study so that preliminary results would contribute to the design of a succeeding trial.

Examples of specific questions are:

1. Is it necessary to subject patients to toxicity in order to achieve a favorable response and, if so, how much toxicity must they endure?
2. Does multimodal therapy enhance or reduce toxicity compared to unimodal therapy, eg, does radiotherapy enhance Adriamycin-related cardiomyopathy?
3. Does combination chemotherapy produce an antagonistic, synergistic, or additive effect on toxicity, eg, does L-asparaginase reduce methotrexate toxicity?

The first question is fundamental to the evaluation of toxicity in oncology trials. Several investigators have approached this problem. Bross et al (9,10) investigated whether host toxicity was necessary to effect a positive response in five phase II studies. For each study they compared response in patients with and without toxic reactions and concluded that toxicity was not necessary (4). They did not, however, answer the fundamental question: does toxicity experience affect response? This simple comparison ignores possible selection factors. In addition, the hypothesis being tested might be self-fulfilling because those patients who continue longer in the study are more likely to experience toxicity than are those who drop out early because of treatment failure.

The first question is really asking the following: What is the effect on response of a treatment policy that intensifies a given therapeutic regimen to generate toxicity in the patient? Creech et al (11) approached this question by a phase II study of low-dose CMF (cyclophosphamide,

[3]Von Hoff D. Personal communication.

methotrexate, and 5-fluorouracil) in patients with previously treated metastatic breast cancer. This group had less gastrointestinal toxicity and myelosuppression, but had a response comparable to that of historical controls treated with high-dose CMF. This nonrandomized study had several inferential problems, among them the fact that patients in the two dosage groups were not comparable with respect to prognostic factors.

Investigation of the first question should be through design in a randomized study. St. Jude Hospital conducted a randomized study of maintenance in pediatric acute leukemia (12). All patients were treated with the same four-drug combination, but were randomized to receive full and half doses. The full-dose group had longer survival time, but greater incidence of severe hematologic toxicity, than the half-dose group.

Bross et al (13) presented a theoretic logistic dose-response model which made various assumptions about the relationship of doses required to produce toxicity and positive response. They demonstrated that sliding-dose protocols, ie, those that titrate dose to produce a given level of toxicity for each patient, appeared to produce the same response, but with lower toxicity than fixed-dose protocols. Currently, the Southwest Oncology Group is conducting a sliding-dose study in pediatric acute leukemia which will estimate the effect of sliding dose on response.

It should be noted that ascertainment bias for toxicity can be introduced into a phase III study through differential frequency of patient contact between treatment groups. For example, if one treatment group receives radiotherapy alone and the other receives radiotherapy plus chemotherapy, patient toxicity data might be collected with greater frequency in the latter group due to more frequent patient contact with chemotherapy. Protocols must specify equal frequency of toxicity data collection in all treatment groups.

## TREATMENT EVALUATION

At the end of a study, each treatment must be evaluated. Table 4 shows a response versus toxicity comparison for the full-dose and half-dose data from St. Jude's (12). A wbc count of $< 2000/mm^3$ and hospitalization for infection are taken as a summary of unacceptable hematologic toxicity. Is the loss in toxicity justified by the gain of survival afforded by the full dose? Most investigators would answer in the affirmative, but the study does point to the need for improved methods of infection control for children with leukemia in remission receiving chemotherapy. This experience illustrates the role of toxicity data in final evaluation of treatment. The purpose is not to rule out effective treatments, but to suggest further improvements in treatment regimens to increase quality of life.

TABLE 4.—Response vs toxicity: 1971 St. Jude Hospital study

| | Treatment group | |
|---|---|---|
| | Full-dose | Half-dose |
| No. of patients | 21 | 21 |
| Median survival (mos) | 34 | 24 |
| Toxicity | | |
| % of patients with— | | |
| Leukopenia (wbcs $< 2000/mm^3$) | 71.4 | 19.0 |
| Hospitalization and/or death from toxicity | 61.9 | 14.3 |

Byar and Corle (14), using data from a clinical trial for prostatic cancer, have recently applied an analysis of treatment—covariate interaction—to the problem of deciding which treatment is best for which kinds of patients. For each patient type in which a superior treatment emerged, the benefit of the one treatment over the other is quantified by the estimated gain in survival. A summary of frequency of unacceptable toxicity could be estimated for each treatment for each of these patient types. This estimate would not influence choice of treatment, but would indicate how far we must go to find the "optimal" treatment for each patient.

## LATE EFFECTS OF THERAPY

Finney (15) has outlined the statistical issues involved in the analysis of late effects of cancer therapy. Attention must be given to improving methods for estimating relative risk of late effects and for minimizing ascertainment bias. In the Southwest Oncology Group, a computer-generated "off-study alive list" is sent to each investigator periodically to remind him/her that survival, performance status, and late effects data should be submitted for each of his/her patients in the long-term followup enumerated on this list. The three aspects of long-term followup—survival, performance status, and late effects—should not be considered mutually exclusive; poor performance status or death might be a consequence of late effects.

Late effects analysis is another area in which intergroup studies would be appropriate. Indeed, the recently established Late Effects Study Group is an international cooperative program that pools data from ten large cancer centers to facilitate late effects research in childhood cancer (16).

## SUMMARY AND CONCLUSIONS

As cancer treatments become more and more effective, increasing attention will be paid to cogent descriptions of toxicity. Investigators must (*a*) strive toward the definition of an appropriate toxicity data set; (*b*) develop appropriate interim and final-data analysis content and

methods; (*c*) develop useful toxicity indices and an understanding of risk factors; (*d*) devise improved methods of building toxicity questions into protocol design; (*e*) develop a mechanism for pooled intergroup analysis; and finally, (*f*) develop methods to incorporate prior information such as that from earlier phase II trials into inference.

The development of this new methodology will only come about if first, the importance of toxicity evaluation in phase III trials is recognized; second, current methods are evaluated and re-evaluated; and third, clinicians recognize their responsibility to motivate and guide statistical office personnel to pursue these various areas of toxicity evaluation.

Finally, a word to the statisticians. Toxicity data are highly complex. They represent non-Markovian repeated measures of both discrete and continuous multivariate random variables with considerable measurement error, considerable potential to be missing or censored, and considerable potential for the existence of prior information. Statistical methods which project to scalar variables, seek simple structure, incorporate prior information, and permit inference in the frequency and time domains frequently answer questions of no clinical significance and/or require assumptions that do not exist in practice. After statisticians become more familiar with the vital clinical questions in toxicity, their greatest challenge will be to develop or improve statistical methods to answer them.

## REFERENCES


1. Gehan EA. Estimating survival functions from the life table. J Chronic Dis 21:629–644, 1969.
2. Breslow N, and Zandstra R. A note on the relationship between bone marrow lymphocytosis and remission duration in acute leukemia. Blood 36:246–249, 1970.
3. Morrison DF. Multivariate Statistical Methods. New York, NY, McGraw-Hill Co, 1967.
4. Vaitkevicius VK, Brennan MJ, Beckett VL, et al. Clinical evaluation of cancer chemotherapy with 5-fluorouracil. Cancer 14:131–151, 1961.
5. Von Hoff DD, Rozencweig M, Layard M, et al. Daunomycin-induced cardiotoxicity in children and adults: a review of 110 cases. Am J Med 62:200–208, 1977.
6. Turnbull BW, Brown BW, and Hu M. Survivorship analysis of heart transplant data. J Am Stat Assoc 69:74–80, 1974.
7. Mantel N, and Byar DP. Evaluation of response time data involving transient states: an illustration using heart transplant data. J Am Stat Assoc 69:81–86, 1974.
8. Crowley J, and Hu M. Covariance analysis of heart transplant survival data. J Am Stat Assoc 72:27–36, 1977.
9. Bross IDJ, Rimm AA, Slack NH, et al. Is toxicity really necessary? I. The question. Cancer 19:1780–1784, 1966.
10. Bross IDJ, Rimm AA, Slack NH, et al. Is toxicity really necessary? II. Source and analysis of data. Cancer 19:1785–1795, 1966.
11. Creech RH, Catalano RB, Mastrangelo MJ, et al. An effective low dose intermittent cyclophosphamide, methotrexate and 5-FU treatment regimen for metastatic breast cancer. Cancer 35:1101–1107, 1975.
12. Pinkel D, Hernandez D, Borella L, et al. Drug dosage and remission duration in childhood lymphocytic leukemia. Cancer 27:247–256, 1971.
13. Bross IDJ, Rimm AA, Slack NH, et al. Is toxicity really necessary? III. Theoretical aspects. Cancer 19:1796–1804, 1966.
14. Byar DP, and Corle DK. Selecting optimal treatment in clinical trials using covariate information. J Chronic Dis 30:445–459, 1977.
15. Finney DJ. Statistical problems in detecting the late consequences of therapy. Cancer 37(suppl):1226–1232, 1976.
16. D'Angio GJ, Meadows A, Mike V, et al. Decreased risk of radiation associated second malignant neoplasms in actinomycin D treated patients. Cancer 37(suppl):1177–1185, 1976.

# DISCUSSION II: Critical Definitions and Criteria for Evaluation*

**Dr. Byar:** Yesterday we discussed patient heterogeneity and the problem this poses when evaluating survival. Today the emphasis seems to have been on heterogeneity of treatment and problems of reproducibility as it applies to methods of evaluation.

Philosophically everyone would agree that the way to proceed would be to control everything that matters and ignore everything that does not. The difficulty, of course, is how to distinguish between these two categories. If we do not know whether or not some variable is important, then it seems prudent to try to control each variable, and this suggests we should standardize nearly everything. However, I believe that our clinical trials and the lives of the statisticians and clinicians involved in them would be greatly simplified if we would expend more effort in determining which factors really matter and less on standardization for standardization's sake.

Recently, I reviewed a large mass of data on estrogen receptors in patients with advanced breast cancer treated by hormonal therapy; these data were gathered at eight medical centers in five countries. Nothing was standardized, not even the measurement of the estrogen receptor, and certainly not the hormonal treatment. The striking thing was that the importance of the receptor measurements in predicting response to endocrine therapy was so strong that the absence of standardization did not seem to make much difference.

The paper by Dr. DeCosse on standardization in surgery was extremely interesting, although it contained a strange dichotomy. The paper describes the number of lymph nodes found in mastectomy specimens, and the summary and discussion dwell on the need for standardization in surgery. Admittedly, these two subjects may be connected, but his paper asserts that there is no proof that the presence of more or fewer nodes is of any prognostic significance. Nevertheless, it is probably desirable to try for some standardization. Concerning the results that Dr. DeCosse presented, two questions puzzle me. First, pathologists differ markedly in the numbers of nodes they find in specimens, but are we invited to believe that all of the pathologists try equally hard? Second, is it not possible that women may differ in the number of nodes they have? Does every woman have, say, exactly 25 or exactly 40 axillary nodes?

The National Cancer Institute has just funded a study of lymph nodes in surgical specimens in women with stage I breast cancer, because it has been estimated that about 25% of the tumors recur by 10 years. It is postulated that by sectioning lymph nodes more closely, one might find micrometastases, not detected at the first examination, which might account for the recurrences.

Dr. Glicksman continued the discussion of standardization in radiotherapy, in pathology, and in the administration of chemotherapy. I agree that central pathology review is essential in every trial where it is feasible. Data for three or four cancer sites I have studied in detail showed that the pathology of the tumor has been the single most important prognostic variable. It follows that if the choice of treatment depends on the pathologic assessment, central review is essential before randomization. However, if tumor pathology is only to be used in treatment comparisons, then all the slides may be sent later for central review, and a stratified analysis may be used to correct for any differences noted subsequently. Dr. Glicksman described an innovative use of the computer in checking whether people were following the protocol, giving the treatment they were supposed to, and ensuring that patients were seen when scheduled.

I was most interested in Dr. Glicksman's description of an attempt to determine whether the mode of administration of radiotherapy to the central nervous system was an important prognostic factor. Of 45 patients chosen randomly from a large study, nine who had correct fields and dosimetry had no disease recurrence in the central nervous system—but there were three recurrences among the 36 who did not have correct fields and dosimetry. If among those 36, 30 were patients for whom the information was incomplete (we do not know that the radiation was incorrectly administered), this result would be statistically significant. Here the result is not significant, but studies of this sort deserve our support because they are asking a vital question: does it matter if you deviate from the protocol?

Dr. Leventhal provided a valuable review of a variety of current approaches to immunotherapy. The possible complications of immunotherapy are daunting: fever, abscesses, activation of old tuberculosis, increased tumor growth, possible development of isoniazid-resistant strains of BCG, infectious hepatitis, serum sickness, anaphylactic shock, and death. Unfortunately, it appears that we are using these therapies before we really know what we are doing. Much of the literature on immunotherapy deals with animal tumor systems, and that seems to be where the greatest success is being achieved.

In most of these animal studies the tumors, rather than arising spontaneously, have been deliberately induced by

*Reprint requests to: Dr. David P. Byar, Biometry Branch, National Cancer Institute, Landow Bldg, Rm C509, National Institutes of Health, Bethesda, MD 20205.

viruses, radiation, or high doses of chemicals over a short period of time. One author has suggested that much of the field of tumor immunology may be a seductive artifact and that the phenomena being observed relate chiefly to how the tumors were induced. Despite this disquieting caveat and the serious concern that for many human tumors immunotherapy may still be premature, this modality continues to provide great hope for the future, if only because we know so little about its potentialities.

Dr. Schmale presented a general discussion of the importance of psychosocial factors in cancer. He pointed out that anyone who wants to study these problems must surmount many stumbling blocks. One of the worst is that everybody approaches the subject differently, so there is no replication or building on past experience. Everyone seeks to come up with new ideas rather than extend or further develop previous ones. Dr. Schmale believes that the first challenge is to identify the frequency and duration of the stresses and strains caused by cancer at every stage, from screening to diagnosis to treatment and recovery. Only after this has been accomplished, he says, will it be possible to assess intervention options. Some work done to date has indicated surprising findings. For example, one study showed a negative relationship between the Karnofsky rating (the way the doctor assesses patients) and a self report of well-being (the way the patients look at themselves). If I understand this finding correctly, it suggests that the more healthy patients were quite dissatisfied, whereas the sicker patients seemed resigned to their fate.

The paper by Dr. Feldstein recommended specific methods for studying agreement between observers when rating patients on integer scales. I found this paper disturbing despite my admiration for his algebraic ingenuity. Dr. Feldstein assumes that the rating scale consists of the equal-spaced integers, but in many medical problems it would be preferable to assume simply that the categories on the scale are ordered without the additional assumption of equal-spacing. For example, suppose we rate psychologic functioning on a scale from 1 to 5. We have no assurance that the difference between ratings 4 and 5 is the same as that between 2 and 3. Nevertheless, the ordering of scales might be useful, and we would like to know how well observers agree in assigning scores to patients. After assuming that the scores are on the equal-spaced integers, Dr. Feldstein explores statistical methods based on the Poisson and geometric distributions which are defined on those numbers. However, as he himself points out, there is no way of testing the validity of these distributional assumptions in the absence of replication, which is rare in such studies. The most serious problem, however, is that Dr. Feldstein's approach ignores the specific items rated and uses only the total scores assigned for all items by each of the judges. Although this results in considerable mathematical simplicity, it fails to give a satisfactory answer to the problem. In effect, Dr. Feldstein is simply estimating whether some judges tend to rate higher than others, rather than how well they agree on the individual items. The fallacy of this approach can perhaps best be understood by a simple example. Suppose two judges are rating five items on a scale going from 1 to 5. Suppose the first judge rates the five items in the order 1, 2, 3, 4, 5 while the second judge rates them 5, 4, 3, 2, 1. All four tests he proposes give chi-square values of 0 which imply perfect agreement between the judges, but I think most people would look at the item-specific scores rather than the test statistics and conclude that the agreement is about as bad as it could be, despite the fact that both judges happen to agree on item 3. Because of the odd behavior of his statistics, I would not recommend their use in practice unless we simply want to know whether some judges tend to rate higher than others.

Dr. Morrow discussed a technique referred to as "multitrait-multimethod matrix analysis," as an approach to the validation of psychosocial measures in oncology research. This technique is essentially an examination of correlations between presumably ordinal scales obtained for different traits or constructs (eg, hope, anxiety, sorrow) by different methods such as verbal report, observer report, or self report. The goal of examining these correlations is to validate a measuring instrument in the absence of any generally accepted objective standard. It is argued that there is a fundamental difference between social and physical sciences in the interpretation of validity. I wonder to what extent this is true. Do we really believe that in the physical sciences we have some perfect outside objective measurement against which to test a new instrument? Even mass, time, and distance are constructs. It is simply that they are better defined than those in the psychosocial field. The really interesting question would be whether methods used in other branches of science could also be used in validating psychosocial measurements or whether because of the necessarily greater imprecision of these measurements different techniques are required. Dr. Morrow briefly presented the characteristics of five other approaches to this problem, but all of them appear to be more or less arbitrary and somewhat subjective. It is clear that more work is needed in this difficult area, and I congratulate him for trying to take a systematic approach to the problem.

Dr. Vietti discussed toxicity scales used by different study groups. One of the problems in assessing toxicity is to determine whether it is due to the cancer or to the treatment. Most people believe they know the answer to this question in any particular case, but in one randomized trial comparing chemotherapy and placebo, significant marrow depression was found in the placebo group. If this study had not had a control group, such toxicity would surely have been attributed to the treatment. Dr. Vietti asked a most important question, namely, whether

toxicity is necessary to achieve an adequate chemotherapeutic response. She presented data suggesting that toxicity may not be necessary. She also presented scales ranging from 0 to 4 which have been used to rate many different kinds of toxicity, and noted that these scales are in general use by most of the large cooperative groups. She has done a useful piece of work in describing how the various groups differ in their use of these scales, and I believe a study of this paper may help reconcile some of the discordant results one finds in the literature.

Dr. Herson discussed statistical approaches to evaluation of toxicity. Undoubtedly, many people have been killed by treatments for cancer. It has been suggested that homeopathy arose not because its proponents thought that it was effective, but rather because they were disgusted with the toxicity produced by purging and bloodletting. More recently, some people have suggested that the hidden agenda of Laetrile proponents is their refusal to accept big medicine, expensive treatment, and horrible toxicity. Papers on the treatment of cancer usually deal with toxicity in a superficial manner, simply presenting tables showing how many people have had nausea, vomiting, etc, at any time during the course of treatment. It is natural that until now it should have been that way because physicians' interest has been in trying to eradicate the cancer. The patient, however, is much more aware of the toxicity. Simple summaries do not take into account the duration of the toxicity or how badly it affects the patient's life.

Dr. Herson urges us to use our best statistical insights to determine how to describe and analyze data concerning toxic reactions. He presented some statistical techniques which could be used to study these phenomena and pointed out the many pitfalls which one most avoid if one is going to make sense of toxicity data. The examples included problems of ascertainment bias, regression to the mean, assessment of entry into transient states, problems of missing data, and effects of outliers. I think Dr. Herson has posed the important questions that any statistical approach to the evaluation of toxicity should answer, and I compliment him on a thoughtful and stimulating paper.

# The Research Logic of Radiation Oncology in Combined Modality Therapy

**Philip Rubin* and Omar M. Salazar[1]**

## SUMMARY

**The integration of radiation, chemotherapy, and surgery depends on the delineation of failure patterns for each specific cancer. By identifying the failure pathways it is possible to design therapeutic strategies for each pattern. The main issues focus upon selecting the best therapeutic sequence and combination that can be employed first to reduce the tumor to subclinical stages, but ultimately to achieve a total cure. Although radiation therapy and chemotherapy can be given simultaneously or sequentially the variety of drugs now available and the many dose schedules for radiation therapy and chemotherapy make the possible permutations almost infinite. It is for this reason that the drug-radiation therapy interactions should be better understood, and that is why new terminology is being developed to describe the different types of interactions that may occur. With the innovations in radiation therapy (systemic half-body irradiation, radiosensitizers, and hyperthermia), there are now more potential combinations of standard modalities and new investigative techniques than can be explored clinically. For this reason, experimental models may be able to offer a better scientific basis for combining modes, but even in the laboratory too many probabilities exist. The need for a cooperative effort (both in laboratory and clinical investigations) for combined modality treatment approaches to cancer has become a mandate for oncology protocol design in this multimodal era.**

[Cancer Treat Rep 64:473–480, 1980]

## RESEARCH LOGIC BASED ON PATTERNS OF FAILURE

Because the first site of relapse often determines the therapeutic strategy, the integration of radiation therapy, chemotherapy, and surgery depends on the delineation of the failure patterns of each specific cancer. In most patients with cancer, treatment fails for three reasons: (*a*) inability to control the primary tumor (T failure); (*b*) nodal involvement (N failure); and (*c*) dissemination to distant visceral sites (M failure). Thus, as shown in the Venn diagram (fig 1), failure may occur from one of these categories or from a combination of factors: a total of seven failure pathways.

[1]Division of Radiation Oncology, University of Rochester, NY.
**Reprint requests to:* Philip Rubin, MD, Division of Radiation Oncology, University of Rochester Cancer Center, Rochester, NY 14642.

### Seven Failure Pathways

*Primary T failures.*—Inability to control primary tumors, particularly in locally advanced stages with invasion of contiguous structures and viscera, remains a significant obstacle to cancer cure. Suit (1) estimated the failure rate of advanced cancers at 12 specific sites, and demonstrated that more than one third of all cancer deaths are, in fact, due to local failures. Some of these tumors are not readily eradicable, such as cancers of the oral cavity, oropharynx, and nasopharynx (T3, T4, N2, and N3); gynecologic cancers of the cervix (stages III and IV), ovary (stage IIB), and vulva; and large advanced genitourinary cancers of the bladder (stage $B_2$ and C) and prostate (stage C). Gastrointestinal tumors frequently begin insidiously and can reach advanced stages in the stomach, colon, and inaccessible sites such as the pancreas. Soft tissue sarcomas are often locally recurrent and invasive before metastasizing. The problem of locally

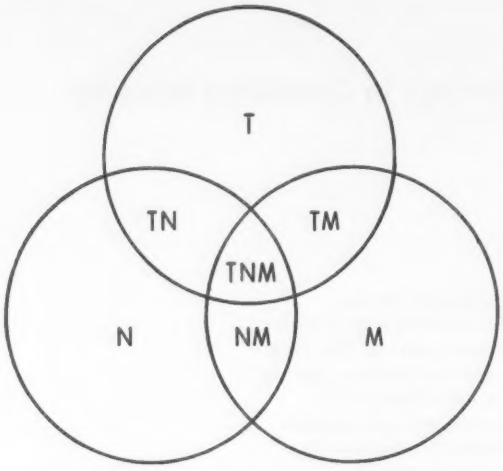



FIGURE 1.—Venn diagram of the seven failure pathways for T, N, and M and their combinations (*from* ref 2).

advanced disease is well-illustrated by malignant brain tumors; for example, glioblastoma multiforme, medulloblastoma, and mixed gliomas are locally invasive and rarely disseminate outside the central nervous system.

*Nodal (N) failures.*—Nodal failures can occur at major node-bearing regions that drain the site of invasive cancers. Metastatic lymph nodes, which can be difficult to control, include cervical nodes in head and neck cancer, axillary and supraclavicular nodes in breast cancer, mediastinal nodes in esophageal and lung cancers, para-aortic and pelvic nodes in gynecologic and genitourinary cancers, and mesenteric nodes in gastrointestinal cancers. As primary tumors are controlled, nodal involvement becomes a source of major failure and requires extended-field radiation at many sites. The introduction of total nodal irradiation techniques for Hodgkin's and non-Hodgkin's lymphomas has focused attention on the control of nodal extensions due to re-seed mechanisms, whether the failure patterns are random or nonrandom.

*Metastatic (M) failures.*—Treatment of the majority of cancer patients "fails" as a result of disseminated disease. The most common cancers, those of breast and lung, are aggressive and metastasize in the early stages of growth, often before clinical detection. Undifferentiated and anaplastic cancers generally metastasize aggressively; melanomas, sarcomas of the bone and soft tissue, chondrosarcomas, and ovarian cancers also may be widespread at the time of diagnosis. Although metastases affect more than one organ system, recognition of the first metastatic site may suggest elective treatment to specific organs.

*Combinations of T, N, and M.*—Four additional failure pathways (TN, NM, TM, and TNM) represent combinations of T, N, and M compartments. In many advanced cancers of the breast, lung, head and neck, and gynecologic sites (T3, T4, N2, and N3 groupings in stages III and IV), both primary and nodal involvement are difficult to resect completely or to eradicate by irradiation. For example, in testicular and ovarian tumors, treatment may achieve local control of the primary, but the main failure pathways are in nodal and metastatic spread. Small cell cancer, other types of lung cancer, advanced gastric and rectal cancers, and pancreatic cancer can and often do fail in all three compartments. Although the T and N categories may be the overt areas of failure, better local control is certain to mask occult metastases.

## CANCER TREATMENT STRATEGIES

Once the failure pathways have been identified (2), it is possible to design specific therapeutic strategies for each pattern. Failure may be due to persistent disease that has not responded favorably to conventional therapy, or to recurrent disease that re-appears after eradication, perhaps in a new compartment. These treatment strategies, briefly tabulated according to failure patterns in table 1, emphasize the potential of combined radiation therapy and chemotherapy, assuming that complete surgical resection has been done when possible and where indicated. For primary (T) and/or regional (N) failures, tumor persistence due to unresectability is generally managed by preoperative or postoperative irradiation. Post-irradiation conservation surgery, a relatively new concept, has been recommended for metastatic neck nodes (3). The ability of chemotherapy to complement irradiation has been demonstrated in advanced head and neck cancers, brain tumors, and gynecologic and testicular cancers. Although radiation therapy is used widely to treat microscopic residua in the surgical bed, chemotherapy may prove to be as effective in selected sites, such as the breast (4). The treatment of patients with high risk of dissemination, and of those who relapse locally and regionally, calls for the prompt integration of modalities, since control of overt nodal recurrence is very difficult. Although irradiation can control suspected microscopic lymph node deposits with moderate doses (4500–5000 rads), oncologists are challenged to devise an effective combined treatment to treat distant metastases without increasing the undesirable toxic side effects.

Overt metastases (M failures) are the most difficult to control with current methods of treatment. High-risk patients should be treated electively, since the majority of those who eventually die of metastases are apparently free of metastatic disease when the primary cancer is detected. The most successful strategy has been to eradicate occult micrometastases with multiagent adjuvant

TABLE 1.—Therapeutic strategy according to failure pattern

| | Failure pattern | Therapeutic strategy—* T | N | M |
|---|---|---|---|---|
| T | Persistence† | X + C | | |
| | Recurrence‡ | X or C | | |
| N | Persistence | | X + C | |
| | Recurrence | | X | |
| M | Persistence | | | C + X |
| | Recurrence | | | C ± X |
| TN | Persistence | X + C | X + C | C |
| | Recurrence | X ± C | X | C |
| TM | Persistence | X + C | | C ± X |
| | Recurrence | X | | C |
| NM | Persistence | | X + C | |
| | Recurrence | | X | C |
| TNM | Persistence | X + C | X + C | C + X |
| | Recurrence | ± X ± C | ± X ± C | C |

*X = radiation therapy preferred; C = chemotherapy preferred; X + C = combination usually used; X or C = either is used; X ± C, C ± X, and ± X ± C = with or without a particular modality.

†Persistence = tumor persistence due to the inability to remove the tumor surgically or ablate it with conventional irradiation.

‡Recurrence = relapse in a given compartment after complete excision; reliance on other treatment modes when failure occurs.

chemotherapy, as in children with Wilms' tumor (5), embryonal rhabdomyosarcoma (6), and Ewing's sarcoma (7,8). We assume that most cancer patients have micrometastases at the time of presentation, but these occult deposits are undetectable with present methods of diagnosis. In order to make effective use of the ability of radiation therapy to sterilize occult metastases in specific sites, we need a careful mapping of spread patterns; this implies knowledge of regional and distant possibilities of dissemination. The calculations of Herring and Compton (9), based on Lionel Cohen's formulation, show that doses in the range of 2000–3000 rads can eradicate cell deposits numbering from $10^2$ to $10^5$ (fig 2). It is possible to irradiate an entire organ, such as the lung or the brain, electively. Because of the blood-brain barrier and the inability of chemical agents to reach tumor cells in the central nervous system, irradiation of sanctuary sites in the brain has been advocated for other highly metastatic tumors; this technique has been successful in the treatment of acute lymphatic leukemia (10) and is applied in an analogous fashion to circulating cancer cells.

In summary, new therapeutic strategies for cancer control should be based on the failure pattern of a given cancer at each stage of presentation (table 1). In addition, all modalities should be optimally combined to achieve the necessary control in the three categories of T, N, and M. Combined irradiation and chemotherapy will augment each other in attempts to control advanced disease in the primary (T) or nodal (N) compartments; in a supplementary or complementary manner, combined treatments should be used to control suspected occult micrometastases. Patients with a high risk of recurrence are the best subjects for clinical trials and protocols investigating the elective addition of treatments.

## THERAPEUTIC SEQUENCE AND COMBINATIONS

Empirical rather than scientific evidence has dominated the selection of those chemotherapeutic agents used in combination with irradiation; success to date has been promising, but limited. At first, selection of combined treatments was based on attempts to understand the mechanisms of drug and radiation actions. However, such choices did not produce the expected differential effects, that is, cell-kill in tumors and a sparing of normal tissues. Often it seemed that the choice of drug and radiation dose reflected chiefly the special interests of an investigator, group, or institution in a particular drug(s) or radiotherapy schedule. Finally, the conflicting statements and the variety of formulations proliferating in the literature give ample evidence of lack of coordination between radiation and chemotherapy.

### Choice of Sequence

Chemotherapy and radiation therapy can be given in three basic combinations: simultaneously, and sequentially in either order, but the variety of drugs now available and the many dose schedules for radiation therapy and chemotherapy make the possible permutations almost infinite. The applications and sequencing of chemotherapy and radiation therapy have varied from anatomic site to organ site, with chemotherapy being used

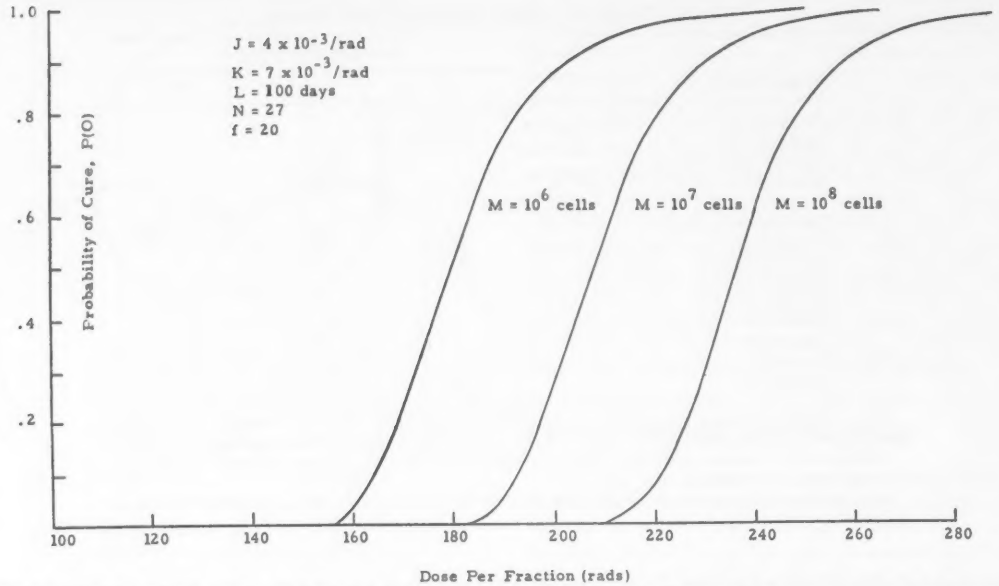



FIGURE 2.—The probability for cure of a tumor is a function of the radiation case for a given set of parameters, and is dependent on the exact values chosen for those parameters. However, the general sigmoid curve of probability of cure vs radiation dose is not sensitive to the particular values of the parameters. In this example, the radiation therapy program contains 20 fractions for an epidermoid tumor consisting of $10^6$, $10^7$, and $10^8$ cells. J, K = radiosensitivity constants for single- and multiple-target components, respectively; N = extrapolation number; L = regeneration or growth constant; F = No. of fractions; M = No. of cells (*from* ref 9).

before, during, and after radiation therapy; there has been little or no consistency in approach. Guides to treatment seem to give more consideration to toxicity and the limits of patient tolerance than to tumor control.

For locally advanced or regional nodal disease, chemotherapy used *before* irradiation is expected to produce maximal tumor regression, thereby allowing for a more effective full dose of radiation or a reduction in the effective dose or treatment volume. It is also hoped that chemotherapy will "sterilize" micrometastases and will limit further metastatic spread while the T and N compartments are being treated with conventional surgery or radiation therapy. Occasionally a chemical agent is administered as a radiosensitizer to "prime" the tumor for radiation. Experience in the treatment of acute lymphatic leukemia (10) has demonstrated that chemotherapeutic induction eliminated circulating cancer cells, and that radiation can destroy the remaining cells in sanctuary sites.

Chemotherapy is employed *during* radiation therapy on the assumption that it exerts a separate, additive, or interactive effect, but these effects are difficult to measure or assess. Toxicity to normal tissue is the principal guide to dose levels (11). Ideally, radiation therapy and chemotherapy should have different target organ toxicities; this is unfortunately not always the case. Chemotherapy during radiation therapy has produced few major improvements. There has been some measurable gain in survival in weeks (brain cancer and glioblastoma multiforme) and in months (stomach cancer) (12), but the simultaneous use and sequencing of these agents need further exploration to determine optimal dose and time interval selection.

The greatest use of chemotherapy has been *after* irradiation and/or surgery, with the expectation that drugs can eradicate micrometastases and eliminate subclinical deposits in vital viscera.

## Choice of Combination

The integration of effective chemotherapy with local surgery or radiation therapy has improved the cure rates of a few solid tumors in adults. This combined approach recognizes that surgery and radiotherapy are local modalities that can kill tumor cells only where they are applied; they fail to cure many patients, even when they remove all tumor visible to the naked eye or on diagnostic X-ray film. This failure appears to be due to the presence

TABLE 2.—New terminology in combined modality therapy as currently utilized in clinical cancer investigation

| | Type* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Gastrointestinal | Lung, oat cell | Lung, squamous and adenocarcinoma | Breast | Hematologic | Head and neck | Brain | Soft tissue sarcoma |
| Spatial coordination, independent toxicity | + | + | – | + | + | + | – | + |
| Therapeutic additivity, independent toxicity | + | + | – | + | – | + | + | + |
| Temporal consolidation, sequential toxicity | – | + | – | + | + | – | – | – |
| Therapeutic synergism, decreased toxicity, radiosensitization, optimal timing | + | + | – | – | – | + | – | – |
| Systemic radiation therapy, hemibody | + | + | + | + | + | – | – | – |

*+ = actually being conducted or possible application; – = no possible application at this time.

of disseminated microscopic disease foci which remain after excision of the primary tumor, surrounding tissue, and some of the regional lymph nodes.

It has been experimentally established that chemotherapeutic agents have a better chance of eradicating a tumor when the tumor cell population is small, eg, immediately after resection of the primary tumor, or after optimal local irradiation of unresected or partially resected tumors. The inverse relationship between tumor cell population and chemotherapeutic cure has been demonstrated in quantitative terms in mice bearing transplanted L1210 leukemia and in experimental solid tumor systems. Optimal chemotherapy has the potential to eradicate metastatic foci of early disease. Those drug regimens that proved most effective in advanced disease have been the best candidates for a combined approach.

A high degree of cell-kill is necessary to shrink a bulky solid tumor by > 50%, usually the minimum definition of objective regression. If this level of cell-kill could be achieved in the relatively small micrometastatic tumor burden remaining after surgery and/or radiation therapy, perhaps all neoplastic cells could be eradicated. Although this line of reasoning seems to be self-evident, it is contradictory to the facts. That is, it requires only 3–4 logs of cell-kill to induce complete tumor regression, ie, 5–1-cm tumor reduces cell load from $10^{12}$ to $10^{9}$. Subclinical disease actually represents a larger biologic load of cells and is more difficult to eradicate. Thus, the residuum of $10^{8}$ to $10^{0}$ cells may require as much as 8 logs of cell-kill to eradicate the last cancer cells.

Chemotherapy has been most successful in the hematologic cancers, especially lymphocytic leukemia (10) and advanced Hodgkin's disease (13,14). However, these triumphs of cure and long-term disease-free survival have not been achieved in the common solid tumors, which cause most of the cancer deaths. The disparity in results is due chiefly to differences in cell kinetics and the relative accessibility to effective drug concentrations exhibited by cells in fast- and slow-growing tumors. In advanced hematopoietic malignancies, chemotherapy is usually the treatment of choice in induction phases, while in advanced solid tumors, the initial treatment is often surgery and/or radiation therapy, and chemotherapy is employed after failure of local treatment or evidence of dissemination.

## NEW TERMINOLOGY IN COMBINED MODALITY THERAPY

Clinical trials have, up to now, employed presumably effective radiation therapy and chemotherapy. The main issue has been how to effectively combine these agents before, during, or after radiation therapy. It has become evident that toxicity will make necessary certain modifications of treatment. A better understanding of biologic mechanisms and clinical patterns of failure should lead to more rational protocol design for combinations of these two treatment modalities (table 2). The mechanisms for such combinations have been classified by Steel and Peckham (15) under different conceptual categories:

*Spatial coordination.*—When one therapy does not control malignant disease in an anatomic site, the addition of a second treatment that does deal with it may improve results, eg, radiation therapy and/or surgery may destroy the tumor in the local and regional (TN) compartments, and chemotherapy may "sterilize" the occult metastases (M). It is usually necessary to introduce radiation therapy or chemotherapy sequentially and at full therapeutic doses. Thus, the most common type of combined modality approach includes many adjuvant programs such as radiation therapy and/or surgery preceding triple-drug therapy in breast cancer; surgery and radiation therapy

followed by chemotherapy with or without immunotherapy in advanced head and neck cancers; surgery and radiation therapy followed by triple-drug therapy in soft tissue sarcomas; and segmental-field radiation followed by multi-agent chemotherapy in stage I and II Hodgkin's disease or non-Hodgkin's lymphomas.

*Therapeutic additivity.*—The concept of therapeutic additivity invokes the principle of *independent toxicity.* The addition of two therapeutic agents that have separate toxic effects may achieve a greater additional tumor-cell-kill. Different target organs have different toxic endpoints for radiation and for chemotherapy agent(s), but both modes act on the tumor in one of the several compartments, usually in the primary and nodal areas (TN). This therapeutic additivity may require simultaneous and/or concurrent administration of radiation therapy and chemotherapy. This type of additivity is presumed to occur during the administration of radiation therapy and chemotherapy (BCNU) in malignant gliomas where both modes are designed to control the primary tumor (T), assuming different toxic effects for radiation therapy and chemotherapy. The principal toxic effect of chemotherapy is on the bone marrow, whereas the dose-limiting tissue for radiation is the normal brain; however, both agents act on the malignant glioma.

*Temporal consolidation.*—This form of additivity is based on the principle of a sequential toxicity; full-dose chemotherapeutic induction is followed by low-dose radiotherapy or vice versa. This approach postulates that the initial induction treatment is effective and reduces the tumor burden markedly so that it usually achieves complete or at least partial regression ($\geq$ 50% of original tumor) and that subsequent measures can be directed at residual occult disease. In this setting, the sequential mode should be more effective in lower doses, should destroy the remaining tumor, and thus consolidate the gains of the initial treatment. This concept was established in stage III and IV Hodgkin's disease by Prosnitz et al (16) and currently is being more widely tested in other sites and tumor types. This type of approach can be seen in the current management of non-Hodgkin's lymphomas, advanced small cell carcinoma of the lung, testicular metastases to the lung, or pediatric tumors where induction with multiple-drug therapy leads to tumor responses $\geq$ 50%; thereafter, low-dose radiation therapy is employed over areas of previously bulky tumor.

*Therapeutic synergism (or supra-additivity).*—In this situation, the combination of the therapies produces a better result than is obtained by either modality alone. The term synergism or, preferably, supra-additivity, is used to describe this circumstance. Therapeutic synergism postulates that smaller (safer) doses of each mode can be employed, because the combination yields better results than even higher doses of each modality employed independently. This category covers such concepts as radiation sensitization, potentiation, or enhancement. It also embraces radiation therapy optimally timed in relation to chemotherapy or radiosensitizers which can be given in single or fractionated doses. The construction of an isobologram is a valid means of detecting true supra-additivity as distinct from additivity alone (15). The concept of optimal or critical timing of doses and radiation therapy is not new in the laboratory. With increasing frequency in the past 5 years, numerous investigations have indicated that the optimal time for drug administration is related to the time for radiation administration. An illustration of this approach is the use of pulsed BCNU before radiation in brain tumors. Based upon the observations of Wheeler et al (17) that BCNU was most effective given at 15 hours before irradiation, and not immediately before or after in rat 9L brain tumors, a pilot study of man has been instituted in the Eastern Cooperative Oncology Group. An experimental combined modality group has been constituted by the Division of Cancer Treatment (DCT), National Cancer Institute (Bethesda, MD) with a mission to search for optimal timing of combinations of radiation and drugs. A large battery of well-studied tumors are under exploration with many different drugs and radiation schedules. This important concept has yet to be successfully applied. Even the widely used schedule of 5-fluorouracil and irradiation on the first 3 days of treatment for gastrointestinal malignancies does not specify the time interval or sequence.

## INNOVATIONS OF RADIATION THERAPY

*Systemic half-body irradiation.*—This innovative approach for the treatment of metastases encompasses the whole body in large segments. Although radiation therapy has effectively treated focal metastatic disease for palliation, this approach allows for treatment of all metastatic foci. Systemic half-body irradiation was developed to meet the need of treating patients with painful bony metastases. It was introduced by Rider et al (18) during this past decade, but they have only recently reported its effectiveness. In encompassing the body in two large fields, the Toronto group found that single doses of 600, 800, and 1000 rads delivered in one fraction to one segment provided dramatic relief of pain within 24 hours, often lasting until death. Except for a transient acute toxicity (mostly severe nausea and vomiting) and delayed pulmonary toxicity, the procedure is reasonably well-tolerated and there is rapid restoration of bone marrow in the 6–8 weeks between treatment of the upper and lower half of the body. In a more recent RTOG pilot study, Salazar et al (19) treated 40 patients with this technique; the majority of whom received upper half-body irradiation for lung cancer. Reduction in tumor size after 800 rads could be assessed in 13 lung cancer patients. Although the numbers are small, the response rate was 85%; most of the patients had partial responses of 2 months' duration (62%). The complete responders were

isolated cases of all cell types, but the most dramatic responses were in the small cell anaplastic tumors; other histologic types will require coned-down booster fields to the overt and large metastatic foci for complete regression. In addition, these authors have devised a comprehensive premedication program that has reduced to a minimum the acute (transient) toxicity seen with the irradiation of the upper half of the body. Both groups have also advocated the correction of radiation doses for transmission in the lung, and at the present time they note that doses of > 700 rads to the upper half of the body are associated with a 10%–20% incidence of fatal radiation pneumonitis. Following the success obtained in the palliation of pain from widespread metastatic cancer, the systemic half-body irradiation technique is being applied to less advanced tumors with the aim of treating microscopic disease in potential visceral sites. The technique is also being explored in diseases such as disseminated small cell carcinoma of the lung and multiple myeloma, with the aim of consolidating primary chemotherapy gains after partial or complete responses have been obtained from an induction.

*Radiosensitizers.*—The modification of the initial radiochemical events which occur in tissue and in tumors during clinical radiotherapy has significant potential for the improvement of clinical radiotherapy. Two classes of compound exist: *sensitizer* compounds are electron affinic and mimic oxygen, and thus sensitize anoxic cells and provide an oxygen enhancement ratio; and *protective* compounds, in general, contain sulfhydryl groups which compete for free radicals and prevent initial radiochemical lesions. Extensive radiobiologic information now exists which indicates that the hypoxic cell sensitizers are extremely effective in animal tumors. Initial clinical trials with two of the compounds, metronidazole and misonidazole, suggest sufficient sensitization, moderately good tolerance, and a need for more extensive clinical trials. On the other hand, the existing compounds have significant neurologic and gastrointestinal toxicity, and improved compounds are being actively synthesized and evaluated. When an optimal schedule of delivery for these compounds is obtained, the combination of radiosensitizers and chemotherapy together with techniques such as systemic half-body irradiation and/or hyperthermia may lead to innovative and unique combinations for improving the therapeutic ratio.

*Hyperthermia.*—In the last decade hyperthermia has been introduced in an unstructured fashion as a treatment modality in cancer management. Renewed clinical interest is related to the recent improvement in heating equipment and to current thermobiologic and radiobiologic research which indicates a definite improved radiation cell-kill associated with the maintenance of a critical heating temperature level ≥ 41°C–42°C. In theory, the therapeutic ratio should be improved since the thermal tolerance of normal oxygenated cells would not be reduced. Hyperthermia in the 43°C–44°C range will damage both oxygenated and hypoxic normal tissue and tumor cells. The range between 41°C and 42°C may affect the tumor by perhaps enhancing radiation effects on the hypoxic fraction and repair processes. It is essential, therefore, that the normal tissue be maintained below 41°C. The cytotoxicity of heat has been repeatedly demonstrated in a variety of in vivo and in vitro systems. The scientific data generated in more complex organisms such as large animals and man demand accurate thermometry to measure temperature distribution for precise isotherms. The influence of blood flow and physiology is of paramount importance in determining effects in tumors and normal tissue. These thermal effects enhance radiation and also the effects of radiosensitizers and chemotherapeutic agents. Therefore, the combined modality approach is important in achieving the potential usefulness of hyperthermia.

## CONCLUSION

There are other potential combinations of existent standard modalities than can be explored clinically. When we examine both innovations and the large variety of possible radiation-drug combinations that already exist, it is apparent that the development of optimal combinations needs more attention than the continuing search for new approaches.

This is not a readily soluble problem, since the simplistic addition of one mode to another has yielded very limited success. Perhaps experimental models in the laboratory can afford a better scientific basis for combining modes, but even in the laboratory, too many possibilities exist. A search for mechanisms of action at the biochemical and molecular level deserves more emphasis. Clinical investigation which observes unique events in the management of human tumors may provide important avenues for the experimentalist in the laboratory.

## REFERENCES


1. Suit HD. Statement of the Problem Pertaining to the Effect of Dose Fractionation and Total Treatment on the Response of Tissue to X-irradiation (time and dose relationship in radiation biology as applied to radiotherapy). Brookhaven National Laboratory, Springfield, VA: Clearinghouse for Federal and Scientific and Technical Information, National Bureau of Standards, US Department of Commerce, 1969.
2. Rubin P, and Carter SK. Combination radiation therapy and chemotherapy: a logical basis for their clinical use. CA 26:274–292, 1976.
3. Fletcher GH, Jesse RH, and Westbrock KC. Interaction of surgery and irradiation in head and neck cancers. *In* Textbook of Radiology (Fletcher GH, ed). Philadelphia, PA, Lea & Febiger, 1973, pp 166–173.
4. Fisher B, Carbone P, Econonou SG, et al. L-Phenylalanine mustard (L-PAM) in the management of primary breast cancer. A report of early findings. N Engl J Med 292:117–122, 1975.

5. Farber S. Chemotherapy in the treatment of leukemia and Wilms' tumor. JAMA 198:826–836, 1966.
6. Donaldson SS, Castro JR, Wilbur JR, et al. Rhabdomyosarcoma of head and neck in children. Combination treatment by surgery, irradiation and chemotherapy. Cancer 31:26–35, 1973.
7. Rosen G, Wollner N, Tan C, et al. Disease-free survival in children with Ewing's sarcoma treated with radiation therapy and adjuvant four-drug sequential chemotherapy. Cancer 33:384–393, 1974.
8. Johnson RE, and Pomeroy TC. Integrated therapy for Ewing's sarcoma. Am J Roent Rad Ther Nucl Med 114:532–535, 1972.
9. Herring DF, and Compton DMJ. The Degree of Precision Required in the Radiation Dose Delivered in Cancer Radiotherapy. La Jolla, CA, Environmental Medicine, Inc, 1971, p 6.
10. Price RA, and Johnson WW. The central nervous system in childhood leukemia. I. The arachnoid. Cancer 31:520–533, 1973.
11. Carter SK, and Soper WT. Integration of chemotherapy into combined modality treatment of solid tumors. Cancer Treat Rev 1:1–13, 1974.
12. Holbrook MA. Gastric cancer: treatment principles. JAMA 228:1289–1290, 1974.
13. Kaplan HS. Hodgkin's Disease. Cambridge, MA, Harvard University Press, 1972.
14. Rosenberg SA. Hodgkin's disease. *In* Cancer Medicine (Holland JF, and Frei E, eds). Philadelphia, PA, Lea & Febiger, 1973, pp 1276–1302.
15. Steel GG, and Peckham MJ. Exploitable mechanisms in combined radiotherapy-chemotherapy: the concept of additivity. Int J Radiat Oncol Biol Phys, 1978. In press.
16. Prosnitz LR, Farber LR, Fisher JJ, et al. Long term remissions with combined modality therapy for advanced Hodgkin's disease. Cancer 37:2826–2833, 1976.
17. Wheeler KT, Deen DF, Wilson DB, et al. Modification of the in-vivo radiation response of rat 9L brain tumor cells by BCNU. Int J Radiat Oncol Biol Phys 2:79–88, 1977.
18. Rider WD, Fitzpatrick PJ, Fryer CJH, et al. Half-body radiation therapy: an update. In press.
19. Salazar OM, Rubin P, Keller B, et al. Systemic (half-body) radiation therapy: response and toxicity. Int J Rad Oncol Biol Phys. In press, 1978.

# Animal Models of Cancer Immunotherapy: Questions of Relevance [1]

David W. Weiss [2,3,*]

## SUMMARY

**Further progress toward the development of more effective clinical immunotherapy depends on a reasoned formulation of the nature, usages, implications—and pitfalls—of test models in animals and in the test tube; these must be made relevant to themselves and to the conditions of neoplasia in man which they are intended to represent.**

[Cancer Treat Rep 64:481–485, 1980]

The great effort in experimental tumor immunology during the past decade has been based on the premise that the laboratory tumor models utilized serve as veracious prototypes of human neoplasia. Extensive programs of clinical investigation into the possibilities of cancer immunotherapy were in turn initiated (and are being continued today, despite the sparse success attained so far) on the assumption that the accomplishments of immunologic intervention in animal neoplastic diseases are relevant to the clinical situation. However, this relevance is very much in doubt: why has so little been achieved through the varied attempts at immunologic intervention in human malignant disease, when such measures are often effective in experimental animal neoplasia?

First, it might be argued that neoplasia of animals is characterized generically by attributes and determinants that differ fundamentally from those pertaining to human neoplasia. Second, we may have constructed our clinical approaches on the basis of irrelevant and faulty animal models; our laboratory tumor analogs and the inferences accruing from their study may be far less pertinent to clinical cancer in man than we had thought.

The first explanation seems unlikely. A categoric discontinuum between cancer in man and in other species in which effective immunotherapy has been demonstrated appears improbable, both on theoretic grounds and in light of the many de facto similarities in the natural history of neoplastic diseases in man and other vertebrates. Indeed, recent findings suggest the existence of common viral and other carcinogenic factors and of related tumor-associated antigens, even over appreciable phylogenetic distances.

The second explanation, in contrast, appears almost self-evident. It is not animals that as such are inherently false analogs of the human host in neoplasia. Rather, the artifacts that compromise the aptness of animal tumor studies to the circumstances of these diseases in man are generated by the investigator. The artificiality lies with the restricted range of tumors that form the subject of most laboratory investigations, and with the design, conduct, interpretation, and extrapolation to the clinic of the experiments undertaken.

The thesis proposed here is that immunotherapeutic studies with animals *can* be relevant to clinical cancer immunotherapy if their planning, conduct, and interpretation are appropriate. Blind reliance on "automatic" pertinence of laboratory research has come to replace careful analysis of supposition and of data, and this tendency must be reversed. Progress toward the development of telling immunologic modalities of patient treatment now seems heavily dependent on the reformulation of the nature and uses of animal models. It is in the context of this aim that I pose several central questions and hazard a number of comments on the construction of tumor immunotherapy test models.

## PROVISIONS OF RELEVANT ANIMAL MODELS OF CANCER IMMUNOTHERAPY

My first questions are directed at the validity of the preceding argument. The development of adequate, rel-

[1] Some of the considerations discussed in this paper are presented in a different form in another essay on the subject (*see* ref 1). No attempt is made here to review in detail the various aspects of tumor immunology under deliberation. Only a few selected references are included, with accent on the author's personal experience in the field.

[2] Lautenberg Center for General and Tumor Immunology, Hebrew University-Hadassah Medical School, Jerusalem, Israel, and Mt. Sinai School of Medicine, New York, NY.

[3] I wish to express my gratitude to Drs. James Holland and George Bekesi, in whose laboratories I was guest while preparing this article.

**Reprint requests to:* Prof. David W. Weiss, Lautenberg Center for General and Tumor Immunology, Hebrew University-Hadassah Medical School, Jerusalem, Israel.

evant animal tumor models is a demanding and expensive task. Is it, in fact, worth making the effort? Is it possible that extensive animal models are not an absolute requirement in this field? Is it perhaps justifiable to undertake human immunotherapy testing without the construct of large-scale animal models of greater relevance? The reply must be decisively "No!" It is perhaps precisely because the clinical initiative has not relied far more on well-grounded guidelines coming from the laboratory that demonstrable therapeutic benefits have been so elusive.

## Model Constellations

There is no one ideal model of neoplasia, even in one species. What is required is a broad spectrum of models, each focused on distinct questions, each relevant to that prototype of neoplastic disease in the test species and, as far as possible, to the presumably corresponding human tumor. Only by the study of a large number of different experimental tumors, in protocols frequently repeated as each controlled variable is altered and in which similar questions are posed recurrently, can we hope to obtain blocks of "principle information" (1). The crucial determinants of appropriate models are repetition and diversity, overlap of the queries set, control of all possible variables, and constant striving for resemblance to the evolution of the similar neoplastic processes in the autochthonous host. From investigations so constructed, sets of generally applicable axioms may come to light.

Internal model relevance is a core demand, and one that has been given far too little attention. Yet it is a requirement that can be met, if not in all respects in each instance, at least to a considerable extent. Many of the experimentally induced, transplantation-propagated neoplasms that lend themselves well to laboratory investigation have their counterparts in spontaneously occurring tumors in the test species. It is possible to mimic the "natural history" of the corresponding spontaneous growth, and that of the intentionally produced cancer in the primary host. In working with transplanted neoplastic tissue, the breadth of artifact can be reduced by considering in the experimental design the tissue site of origin of the primary tumor, its patterns of invasiveness and metastasis, the strain subline in which it arose, and its kinetics of development vis-à-vis the kinetics of intervention.

An example of the usefulness of *constellations* of model systems is provided by studies undertaken to ascertain the dosage range of nonliving, nonspecific immunomodulators. Determination in vivo and in vitro of the optimum amounts required for immunopotentiation and resistance elevation by the MER tubercle bacillus fraction (2) in several species, in widely varied test models, has given rise to the following policy guidelines.

*a.* There is no evident host body weight-agent dose relationship.

*b.* Small amounts of the adjuvant substance are usually more effective than larger quantities, even where these are well-tolerated.

*c.* Supra-optimal concentrations of the agent are likely to have adverse effects on immunologic and resistance functions.

*d.* The most effective levels fall within similar and narrow ranges, one for application to the intact animal and one for introduction into tissue culture systems, regardless of the test species and of the criteria of stimulation (3,4).

Only such "principle information" as this can serve the critical need for suggestions of conditions and limits for a new treatment when it is first brought from laboratory to clinic. The use of diversified concatenations of models may afford outlines of the prevailing boundaries of effective immunologic therapy and point to the probabilities of notable exceptions; even such modest expectations cannot be entertained for studies with a restricted number of prototypes, no matter how supposedly ideal the isolated model or how impressive the results within the parameters selected by the investigator.

## The Nature of Immunologic Function

Newer insight denotes immunologic reactivity as a labyrinth of positive and negative signals, cellular and humoral, in which a given end result is reached via chains of finely balanced, intermeshing reactions involving numerous helper, effector, and suppressor events. Linked to each other in repeated interactions, components of the immunologic apparatus can substitute one for another, amplify, and abrogate. Antibodies and immunocytes with distinctly diverse end functions are not infrequently the late derivates of common pathways of differentiation and biosynthesis which bifurcate terminally. Seemingly small changes in the complement of reacting factors and in the prevailing circumstances can have categoric effects on the outcome of immunologic response. Our apperception of this overwhelming complexity of immunologic reality has not yet been matched, however, by an analytic dissection of its fabric. Awareness of the multifaceted interplay between reaction pathways forbids us to resort to such terms as "immunologic failure," "immunopotentiation," "immunologic surveillance," and "tumor cell escape" with any pretense of true comprehension. These designations are only grossly descriptive; without further knowledge of the mechanisms, kinetics, conditions, and localities of the events occurring in nature or precipitated by our manipulations, we have little hope of achieving a measure of predictable control. Basic understanding of cause and effect in immunologic functions remains a direct and immediate imperative of the exertions of the cancer immunotherapist. This is not to say that clinical work should now be suspended until we have gained a deeper

apprehension of the underlying immunologic phenomena. The cancer problem is too urgent for such a pause. We must, however, recognize the crudeness of our present suppositions and technical means, and make every effort to found ongoing clinical investigations on emerging immunologic principles.

## Neoplasia as Host-Parasite Relationship

Neoplasia must be viewed as a multidimensional and dynamic interaction among host, tumor cells, and environment; exclusion of large segments of this interaction from our fields of cognition leads to a distorted picture. Emphasis on a given facet of the subject is unavoidable in experimentation, but it must be recognized that what is grasped thereby is only a small part of the whole, changeable as other constituent factors change. Such recognition must be built into overall experimental strategy. Only by a sequential and complementary focusing on events can the fluid links of reaction and interaction that make up the host-parasite relationship of neoplasia be mapped.

Exposition of neoplasia as a host-parasite relationship has been proffered by me previously (5), and only the essential elements of the view are summarized here. Host, neoplastic parasite, and environment are interlinked variables. Environment includes all the extrinsic and intrinsic factors that impinge on the confrontation of host and tumor. The makeup of a host organism is determined by a combination of intrinsic factors, such as genotype, age, sex, metabolic activity, and intercurrent disease, as well as by the elements of its ambient environment. The latter often determine the expression of the former. The anatomic and physiologic individuality of the organism is thus the consequence of a multiplicity of interacting determinants, and is continuously in flux. Moreover, one must bear in mind that the importance of any one determinant, and of the total set, may be very different for general host "health" and reproductive behavior, and for the ability to resist a given challenge—a consideration stressed more than two decades ago by Schneider with respect to host-parasite relationships in the realm of infectious diseases (6).

Phenotypic expression of the deviant genotype of neoplastic parasites is similarly under the aegis of a complex of interrelating environmental circumstances, ambient and of host origin. Moreover, assessment of the influence of total environment on neoplastic properties of transformed cell variants must take into account the *responsiveness* mechanisms of the host. Tumor cell morphology, physiology, and neoplastic behavior may vary as a function of host refractoriness, whether immunologic or nonimmunologic (7,8). The possible influence of host responsiveness must be kept in mind in experiments where tumor cell populations are sheltered or removed from the matrix of host reactivity.

Consideration must be given as well to the nature of the microenvironment of each tumor focus, even within the same host organ, and to the variabilities that go with sequential neoplastic progression and incidental secondary mutation.

It is also mandatory that a philosophy of treatment specifically pertinent to immunologic intervention be developed. Animal models of immunotherapy must not preempt other treatment concepts; only when they are constructed in light of the opportunities and constraints peculiar to the immunologic approach can they possibly provide stepping stones to intelligent immunotherapeutic effort in the clinic.

## Tumor-Associated Transplantation Antigens and the "Central Fallacy" of Tumor Immunology

In its original form, the theory of immune surveillance postulated the existence of unique tumor-specific transplantation antigens (TSTA) as a ubiquitous or at least very frequent property of transformed cells, and ascribed progressive neoplastic disease to immunologic malfunction or to special mechanisms by which tumor cells escape from immunologic control. The common occurrence of TSTA came under question with the finding that at least a portion of the antigenic determinants that characterize the tumor cell surface is also expressed by normal cells, but at different stages in ontogeny or in other tissues. Doubts as to the reality of TSTA were accentuated when it was demonstrated that other tumor-associated antigens are virion, virus-coded, or virus-derepressed in origin, and also occur at times on nontransformed cells, albeit in different amounts or molecular associations.

These observations brought about a shift of emphasis from true tumor-*specific* transplantation antigens (indeed, it no longer seemed possible to ascertain that any tumor cell constituent is wholly specific to the neoplastic state of the cell) to the amorphous category of tumor-*associated* transplantation antigens (TATA) (9,10). Although pronounced deviations from normal in the localization, quantity, and configurations of cell components may indeed endow the deviant clone with "novel" antigenic properties, de facto, it has become questionable whether antigens exist that are *wholly* novel to the neoplastic condition. The cornerstone of the theory of immune surveillance against neoplastic variants further crumbled when it was reported that many spontaneously developing tumors fail to evince antigenic behavior in the primary and in sygeneic hosts (11,12). The demonstrated immunologic propensities of experimental tumors came to be viewed as artifacts arising from the circumstances of their induction in the laboratory, and the axiom of "tumor antigenicity" is now seen by many investigators as the central fallacy of tumor immunology rather than its sustaining principle.

However, this negative evaluation appears to be ex-

cessive. Some or even many spontaneously arising neoplasms may indeed lack immunogenic properties (13,14), but it is still valid to postulate that tumors in nature as well as in the laboratory are distinguished by *antigenic qualities*. The crucial distinction, then, is between antigenicity and immunogenicity. The former designation should refer only to the *potential* of a substance to evoke immunologic responses; the latter describes realization of this potential (13).

The crucial question, then, is by what means can the distinctive but low antigenicity of tumor cells be brought to a level of immunogenicity sufficient to induce transplantation reactions against the neoplastic variants as they are in the host (14)? Clearly, this question cannot be answered with laboratory models of tumors whose degree and kind of antigenicity have little bearing on spontaneous neoplasms in nature; hence, radical change may be demanded in the nature of laboratory tumors used as "standard" test systems. The type as well as the immunogenic vigor of TATA must be considered in the restructuring of relevant animal tumor models. It appears that various kinds of TATA can be located on the membrane of the same cell—embryonic and fetal markers, organ-associated determinants, virus-related antigens, and *perhaps* even TSTA—and that one may be in competition with another for expression. Moreover, in the presence of one category of antigenic determinants, others may be presented to the immunologic apparatus in altered configuration. The role of different groups of TATA, native or modified, in eliciting immune responses is uncertain; this uncertainty might be resolved by systematic analysis of the resistance mounted by syngeneic hosts against neoplasms expressing each class of antigen to known extent. The information thus obtained could contribute appreciably to our ability to excite defensive immunologic reactions against human neoplasms distinguished by certain TATA constellations.

Tumors often change appreciably, their antigenic profile included, after even only a few passages in syngeneic carriers or in tissue culture. Antigenic alterations occurring merely in the process of cryopreservation and thawing have also been reported (15). There is an analogy here to microbial parasites, whose antigenic and pathogenetic properties fluctuate markedly and rapidly upon removal from the infected primary host. Although some of the original characteristics of a neoplastic clone may be restored by re-passage in a host animal, others may be permanently lost from the cell population after as few as one or two passages. Appeals have long been heard for emphasis on work with primary tumors in the autochthonous host, or at least with cells of very early transplant generations in animals of proven isogenicity (16). Unfortunately, most of the effort in tumor immunology has been with ancient lines of "tumor" cells, which may retain little semblance to the neoplastic state in nature. Supposedly syngeneic host animals have often been taken from subcolonies of the strain in which the tumor has arisen, without confirming the maintenance of isogenicity between long-divergent breeding nuclei.

Tumors of the same tissue type as that of many of the long-transplanted neoplasms in common use arise with appreciable frequency in large animal colonies. At least occasional test experiments with such growths can be undertaken in the autochthonous host and in animals of confirmed isogenicity currently produced in that colony. Samples of tumor tissue taken from the original mass or from first-generation transplants can be cryopreserved for future study. Although the procedures entailed may affect TATA properties, freeze preservation under optimal conditions seems less likely to compromise antigenic integrity than repeated animal or tissue culture passage with all the attendant possibilities of selective and adaptive change.

A shift to tumors of domestic animals could much improve the value of animal models. A variety of neoplasms occur in cats, dogs, and several species of farm animals with sufficient frequency to constitute feasible test material. The unavailability of syngeneic hosts is offset by the salient advantages provided by such cancers. The growths arise naturally, and the artificiality of domestication is probably far less than that pertaining to long-inbred lines of laboratory animals. Results from a small number of experiments in the primary tumor host would be inherently more telling than similar data obtained from large experiments in supposedly syngeneic subjects. The large size of domestic animals also permits a variety of auxiliary experimental manipulations; for instance, plasmapheresis before and intercurrent with immunotherapeutic measures (17) may heighten the efficacy of treatment and provide additional analogies to intervention possibilities in the patient.

The emphasis on relevance pertains no less to in vitro models of tumor immunology and immunotherapy than to animal studies. Much attention has been given recently to attempts at in vitro sensitization of cytotoxic effector cells (4,18) and in vitro generation of antibodies (19–21) for passive or adoptive immunotherapy (14). The cytotoxic products thus obtained are usually screened for antitumor cell activity by in vitro tests. Such experimentation is accompanied not only by the artificialities inherent in laboratory studies with intact animals, but also by the large additional distortions from nature involved in all work with isolated tissue and cell systems. Indeed, the results obtained from in vitro assessment of cytotoxic potency and from even such highly contrived in vivo assays as Winn tests often show little parallelism (22,23), and it can be assumed a fortiori that correlation between in vitro measurements and characterization of immunity in the whole animal is generally still more tenuous.

## Standards of Comparison

A word of caution is in order with regard to the standards of comparison invoked when findings from animal (and from in vitro) experiments are applied conceptually to the human situation. It is always formidably difficult in the design and analysis of models to reconcile relative and absolute scales of comparison with the human counterpart. For instance, a one-gram carcinoma transplant in a mouse may represent in absolute mass a pertinent facsimile of early cancer growth in the breast of a woman, but it does not mimic incipient mammary neoplasia in the rodent. In some respects, perhaps, information of comparative value can be gleaned from experiments in which the clinical situation is copied in absolute terms. Even then, however, the investigator must be conscious of the different dimensions of comparison that can be utilized, and he must examine all modalities before deciding which offers the most relevant copy—and these may differ for different phenomena. Good experimental design seems to require that timing and quantification of treatment of model tumors be chosen with a close view to the kinetics of progression of the prototype neoplasms in primary hosts of that test species, ie, again with emphasis on the internal relevance of each model system.

## REFERENCES


1. Weiss DW. Animal models in cancer immunotherapy. *In* Immunotherapy of Human Cancer, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, Houston, TX, 22nd Annual Clinical Conference on Cancer, New York, NY, Raven Press, 1978, pp 101–109.
2. Weiss DW. MER and other mycobacterial fractions in the immunotherapy of cancer. Med Clin North Am 60:473–497, 1976.
3. Ben-Efraim S. Methanol extraction residue: effects and mechanisms of action. Pharmacol Ther [A] 1:383–410, 1977.
4. Kedar E, Nahas F, Unger E, et al. In vitro induction of cell-mediated immunity to murine leukemia cells. III. Effect of methanol extraction residue fraction of BCG on the generation of cytotoxic lymphocytes against leukemia. JNCI 60:1097–1106, 1978.
5. Weiss DW. Neoplastic disease and tumor immunology from the perspective of host-parasite relationships. Natl Cancer Inst Monogr 44:115–122, 1976.
6. Schneider HA. Nutrition and resistance-susceptibility to infection. Am J Trop Med 31:174–182, 1951.
7. Apffel C. Nonimmunological host defenses: a review. Cancer Res 36:1527–1537, 1976.
8. Zinzar SN, Svet-Moldavsky GJ, and Karmanova NV. nonimmune and immune surveillance. I. Growth of tumors and normal fetal tissue grafted into newborn mice. JNCI 57:47–55, 1976.
9. Bentwich Z, Weiss DW, Sulitzeanu D, et al. Antigenic changes on the surface of lymphocytes from patients with chronic lymphocytic leukemia (CLL). Cancer Res 32:1375–1383, 1972.
10. Weiss DW. Immunological intervention in neoplasia. Johns Hopkins Med J (suppl) 3:131–156, 1974.
11. Hewitt HB, Blake ER, and Walder AS. A critique of the evidence for active host defense against cancer, based on personal studies of 27 murine tumors of spontaneous origin. Br J Cancer 33:241–259, 1976.
12. Klein G, and Klein E. Immune surveillance against virus-induced tumors and nonrejectability of spontaneous tumors: contrasting consequences of host versus tumor evolution. Proc Natl Acad Sci USA 74:2121–2125, 1977.
13. Weiss DW. The questionable immunogenicity of certain neoplasms: what then the prospects for immunological intervention in malignant disease? Cancer Immunol Immunother 2:11–19, 1977.
14. Weiss DW. Tumor antigenicity and approaches to tumor immunotherapy—an outline. Curr Top Microbiol Immunol 89:1–83, 1980.
15. Bartlett GC, Katsilas DC, Kreider JW, et al. Immunogenicity of "viable" tumor cells after storage in nitrogen. Cancer Immunol Immunother 2:127–133, 1977.
16. Weiss DW. Immunology of spontaneous tumors. *In* Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (Lecam L, and Neyman J, eds). Berkeley, CA, University of California Press, 1967, pp 657–706.
17. Israël L, and Edelstein R. In vivo and in vitro studies on nonspecific blocking factors of host origin in cancer patients. Role of plasma exchange as an immunotherapeutic modality. Isr J Med Sci 14:105–130, 1978.
18. Kedar E, Nahas F, Schwartzbach M, et al. Generation in vitro of cytotoxic effector cells against syngeneic and allogeneic mouse leukemias and lymphomas. *In* Tumor-Associated Antigens and Their Specific Immune Response (Spreafico F, and Arnon R, eds). New York, NY, Academic Press, 1979, pp 211–235.
19. Globerson A. In vitro approach to development of immune reactivity. Curr Top Microbiol Immunol 75:1–43, 1976.
20. Ben-Efraim S, and Diamantstein T. Mitogenic and adjuvant activity of a methanol extraction residue (MER) of tubercle bacilli on mouse lymphoid cells in vitro. Immunol Commun 4:565–577, 1975.
21. Ben-Efraim S, Sarir C, Dar O, et al. Effects of the MER tubercle bacillus fraction on the production of antibodies in vitro. I. Effect on the primary response. Cell Immunol. 50:314–326, 1980.
22. Miller F, Dexter DL, and Heppner GH. Comparison of microcytotoxicity (MCT), chromium-release test (CRT), and Winn assays in detection of cell mediated immunity (CMI) to mouse mammary tumors. Proc Am Assoc Cancer Res and ASCO 19:38, 1978.
23. Cohen D, Yron I, Grover NB, et al. Comparison of in vivo and in vitro assessments of the immune response of C57B1 mice against a transplantable syngeneic methylcholanthrene-induced fibrosarcoma. Cancer Immunol Immunother, 1980.

# Prognostic Factors in Cancer of the Testis: An Approach to Identification [1]

**Anne Goldman, George Bosl, Kay Johnson, Elwin E. Fraley, and B. J. Kennedy [2,*]**

## SUMMARY

**The clinical experience at the University of Minnesota Hospitals between 1962 and 1977 was reviewed in the hope of defining the biologic features of testicular cancer, knowledge of which is essential to determine the therapeutic needs of each patient. This collaborative effort between clinicians and statisticians has attempted to define important variables in the assessment of risk factors that determine the choice of treatment for testicular cancer. A preliminary analysis of data on the first 78 patients tabulated was performed to test and plan an approach to the final analysis of a larger group of 282 patients. One prognostic factor, staging, is discussed in detail to illustrate the purpose of this exploratory testing. Retrospective analysis, according to stage, of patients who were referred to this cancer center revealed a poorer prognosis than should be anticipated. In reporting the clinical experience of an individual institution, problems of bias are introduced with respect to the nature of the types of tumors studied.**

[Cancer Treat Rep 64:487–496, 1980]

Recent advances in the management of testicular cancer require a re-assessment of the methods of diagnosis and treatment in relation to the known biologic features of this disease. An evaluation of past and present patients can provide data about prognostic factors which will aid in the selection of specific methods of treatment at each stage of this type of cancer. We have undertaken such an analysis of 282 patients using the computer to manage and analyze the data. A preliminary examination was conducted of the data on the first 78 patients using the original histologic classification, for the purpose of testing and planning an approach to the final method of analysis. As an illustration of the testing method, the examination of one factor, staging, as a prognostic indicator is presented.

[1]Supported by research grants from the Minnesota Medical Foundation, the Masonic Hospital Fund, Inc, the American Cancer Society, and the Whirlpool Foundation.

[2]Division of Biometry, School of Public Health (*A. Goldman*) and Section of Medical Oncology (*G. Bosl, K. Johnson,* and *B. J. Kennedy*) and Department of Urology (*E. W. Fraley*), School of Medicine, University of Minnesota, Minneapolis.

**Reprint requests to:* B. J. Kennedy, MD, Box 286, Mayo Memorial Bldg, 420 Delaware St, SE, Minneapolis, MN 55455.

## CLINICAL BACKGROUND

The incidence of cancer of the testis adjusted for age is 3.1/100,000 people. Testicular tumors account for 11.4% of all cancer deaths in the 25–34-year age range. It is also the most common type of cancer for men between the ages of 20 and 34 (1) (table 1).

Although no more than 2000 new patients with nonseminomatous testis cancer are diagnosed yearly in the US, this disease has a catastrophic socioeconomic impact on young men (in early career development), young married couples, and young families. Thus, cure of cancer of the testis would make an important contribution to the economic and cultural aspects of society.

Before 1962, advanced testis cancer was primarily treated with "triple therapy" including actinomycin D, methotrexate, and chlorambucil, but the introduction of mithramycin led to a new approach. Mithramycin proved to be exceedingly toxic, causing hepatic and platelet toxicity, with a mortality rate as high as 24% (2), and for a time was deemed unsuitable for clinical use (3). However, laboratory investigation revealed that this drug has a unique characteristic: a dissociation between the times of toxicity on normal and tumor tissues. With alternate-day administration, mortality from therapy was elimi-

TABLE 1.—Rank order of the most frequent malignancies in 20–34-yr old men (ref 1)

| Rank | Age range in yrs | | |
|---|---|---|---|
| | 20–24 | 25–29 | 30–34 |
| 1 | Testicular cancer | Testicular cancer | Testicular cancer |
| 2 | Hodgkin's disease | Hodgkin's disease | Melanoma |
| 3 | Leukemia | Melanoma | Hodgkin's disease |

nated (4). Mithramycin proved to be uniquely specific for embryonal cell carcinoma of the testis, but had no effect on teratocarcinoma or choriocarcinoma (5). Residual metastatic deposits that did not disappear, such as pulmonary nodules, were surgically removed. In many instances, the nodules that failed to regress proved to be of a histologic type other than embryonal cell carcinoma.

During the next 10 years, the therapy for disseminated testicular cancer was further improved with the introduction of the combination of: vinblastine and bleomycin which produced a 43% remission rate (6); vinblastine and a 5-day infusion of bleomycin which produced a 49% complete remission rate (7); VAB III (vinblastine, actinomycin D, bleomycin, and cisplatin) which produced a 58% complete remission rate (8); and vinblastine, bleomycin, and *cis*-dichlorodiammineplatinum(II) which produced a 70% complete remission rate (9). The last two investigations (8,9) clearly demonstrated that most testicular cancers will undergo some degree of regression following chemotherapy. It also became apparent that only complete remission is meaningful.

Approximately one half of the patients with advanced testicular cancer treated with chemotherapy have a complete remission which lasts for > 2 years. With modern combination therapy, patients who have undergone complete remission of disseminated metastases survive without recurrence, and it can now be asserted that disseminated carcinoma of the testis can be *cured* by chemotherapy.

Since 1972, several factors have changed the management of testicular cancer. Serum biologic markers—alpha-fetoprotein and chorionic gonadotropins which are produced by the tumor or which reflect its presence—have been applied to the diagnosis and monitoring of the disease. New histologic definitions and new morphologic classification systems have been introduced (fig 1). The introduction of the sonogram, computerized tomography, and more extended retroperitoneal lymphadenectomy have improved the staging of this disease (table 2). The need for staging data is emphasized by the statement on testis cancer in the 1977 Manual of the American Joint Committee for Cancer Staging and End Results Reporting that, "No stage grouping is currently recommended" (10).

In 1962, the Section of Medical Oncology of the University of Minnesota Medical School began to investigate the management of advanced carcinoma of the testis; since then, an increasing number of patients have entered these programs (table 3). The increase created problems in management, review, and analysis of the accumulated data. The task of tabulating patient data by hand became overwhelming. Conversion to computer analysis made it possible to re-evaluate the experience of 282 patients with testicular cancer seen at the University of Minnesota Hospitals from 1962 to 1977.

TABLE 2.—Clinical staging of testicular cancer

| Stage | Description |
|---|---|
| I | Confined to the testis |
| II | Spread to retroperitoneal lymph nodes |
| III | Disseminated<br>A. Above the diaphragm<br>B. Extranodal in abdomen |

Some of the variables that must be considered when assessing the management of testicular cancer include the method of staging, histology, biologic markers, methods of treatment, and treatment results. The development of a uniform method of reporting data required: review of the histologic interpretation of the primary tumor, metastases, and autopsy materials; histochemical staining of pathologic materials to update the information that could be obtained from earlier patient experiences; incorporation of new types of information such as the biologic markers; and finally, critical evaluation of the

TABLE 3.—No. of new patients with testicular cancer seen at University of Minnesota Hospitals (Cancer Registry Data)

| 2-yr periods | No. of new patients |
|---|---|
| 1963–1964 | 14 |
| 1965–1966 | 14 |
| 1967–1968 | 25 |
| 1969–1970 | 31 |
| 1971–1972 | 42 |
| 1973–1974 | 52 |
| 1975–1976 | 77 |
| 13-yr total | 255 |

Seminoma
- Classic
- Anaplastic
- Spermatocytic

Embryonal carcinoma

Teratocarcinoma

Teratoma

Choriocarcinoma

Yolk-sac tumor

Other

FIGURE 1.—Histology of testicular cancer.

Date and institution where the event was observed

Physical description
- Symptoms
- Signs
- Tumor location

Method of procedure
- Diagnostic
- Surgery
- Radiotherapy
- Chemotherapy

Histology of tumor

Biologic markers
- Before and after treatment
- As monitors

FIGURE 2.—Groups of data relevant to major events in testicular cancer.

staging systems. This report describes an approach to the statistical and clinical methods used to identify risk factors in the study of testicular cancer by performing a preliminary analysis of the first 78 patients tabulated.

## EXPECTATIONS OF DATA ANALYSIS

Data review should provide specific answers to questions proposed by clinical investigators and should suggest a method of analysis of new patients. This type of study can be directed at the following goals.

*The preparation of data for use in public and professional education.*—Clinical experience indicates that after the onset of a sign or symptom of testicular cancer, patients tend to delay before consulting a physician. A second delay may result from the physician's failure to recognize the presenting complaints and signs as an indication of testicular cancer. Data concerning both kinds of delay can provide material for a program of public education which will make young men aware of testicular cancer and will encourage early visits to physicians. Data concerning factors which contribute to delay by physicians can provide the basis for professional education which will assist earlier diagnosis of cancer of the testis and increase the probability of cure.

*Establishment of risk factors to assess and predict response to therapy and survival.*—An analysis of the multiple variables in testicular cancer and the many factors in treatment should identify those factors which are most significant in altering rates of response and increasing the duration of survival. As has been established with other tumor systems, such as lymphoma (11), multiple myeloma (12), and prostatic cancer (13), the risk factors can be evaluated using multivariate analysis.

*Determination of the role of biologic markers in the diagnosis and monitoring of testis cancer.*—As noted, the introduction of serum alpha-fetoprotein and the beta subunit of human chorionic gonadotropin has promised to improve the diagnosis and management of testicular cancer (14). Correlation of the presence of high levels of these markers with other biologic features of testicular cancer can help define the role of marker monitoring in clinical management.

*Evaluation of therapeutic results.*—The appropriate use of radiotherapy and surgery in early disease and chemotherapy as adjuvant therapy in early disease, or alone in advanced disease, has improved survival. As these techniques improve, the rate of response and duration of survival should increase. The analysis of the results of recent therapies will provide data for comparison with future therapies.

*Development of a staging system.*—Description of the extent of the cancer at the time of the initial diagnosis permits accurate comparison of treatment results. Such staging systems are being developed for all major tumor systems. An objective of the current review is to provide data to define a better staging system for cancer of the testis.

## MATERIALS AND METHODS

The proposed study of 282 patients involved some 300 variables for each. For a study of that size, a surprising amount of the total effort went into the planning of the process of collecting, processing, and analyzing a complete data base. The physical manifestation of this planning is the data collection form. Such forms express all the ideas concerning what variables are available and are relevant to all the central and ancillary scientific questions that may be asked. For the purpose of planning the final analysis of 282 patients, a preliminary examination of the data for the first 78 patients tabulated was performed, and this report is the result of that exploration.

TABLE 4.—Abstraction of data for testicular cancer

| Sequence of events | Date and institution | Physical evaluation | Method and procedure | Pathology histology | Biologic markers |
|---|---|---|---|---|---|
| Symptoms | X | X | | | |
| Physician's visit | X | X | | | |
| Orchiectomy | X | X | X | X | X |
| Lymphadenectomy | X | X | X | X | X |
| Radiotherapy | X | X | X | | X |
| Chemotherapy | X | X | X | | X |
| Response | X | X | X | | X |
| Recurrence | X | X | X | X | X |
| Followup status | X | X | | Autopsy | |

The records of most patients with testicular cancer show a sequence of major events, progressing from first symptoms through orchiectomy and other treatment to final outcome as follows.

True onset (unknown)
First symptoms or trauma
First visit to physician
Orchiectomy (diagnosis)
Lymphadenectomy, retroperitoneal
Radiotherapy
Chemotherapy: adjuvant or therapeutic
Response
Recurrence
Current life status: dead or alive

A specific patient may not experience all stages, and the stages may not always occur in the same order. For example, a patient whose testicular tumor was discovered during a routine physical examination had no "first symptoms." A patient who had a laparotomy for an abdominal tumor and was then discovered to have a tumor in the testis reversed the usual order of surgical procedures (orchiectomy, then lymphadenectomy).

In examining the kinds of data thought to be relevant both for monitoring individual patients and for studying the disease, it becomes evident that a few groups of data are encountered at several points in the sequence of events (fig 2). For example, the histologic classification (15) shown in figure 1 can be applied at the time of orchiectomy, lymphadenectomy, biopsy, and autopsy. All available tissues are being collected and reviewed in order to adapt to this new histologic classification.

The biologic markers have been useful in monitoring the course of the disease (14) but the prognostic importance of some of them is not clear. Unfortunately, large centers began to collect data on markers only recently, and in smaller institutions such collection is not yet routine, so the data are sparse. In addition, assay methods have not been standardized, and in many cases the results are merely noted as "normal" or "abnormal."

To establish the relationship between the sequence of events and the types of data to be abstracted, an outline of the data collection form was prepared (table 4). The data are of two types, retrospective and prospective. From the time a patient is seen at the University of Minnesota, the data are prospective (although abstracted from the charts). However, some patients were not referred to this institution until they required chemotherapy for advanced cancer, while some patients were seen as early as at the time of their orchiectomy. All data collected before the patient was referred to the University Hospitals are retrospective, and so may have the characteristics of such data.

## Analysis

At the University of Minnesota, several years of collaboration between Biometry and Medical Oncology have established the general procedures for identifying and evaluating risk variables. Peto et al (16,17) and Armitage and Gehan (18) described the statistical tools useful for cancer studies. After the data base is set up and edited, it is necessary to determine the distribution of each of the variables and the principal characteristics of the patient populations. (For practical reasons, we do not examine secondary variables unless necessary.) We then examine the relationship between suspected risk variables and survival time. Life tables (usually many of them) are calculated for each group and subgroup according to each risk variable. The relation between each categoric risk variable and response rate is examined by calculating a series of contingency tables comparing patient responses and classification of these variables. When the risk variables are quantitative rather than qualitative, the mean risk in the group of responding patients is compared to that of nonresponders. The final stage of analysis is the study of the multiple interactions of the risk variables by fitting and testing a series of multivariate models of the survival time (18) and multivariate contingency table analysis of the response rates (19).

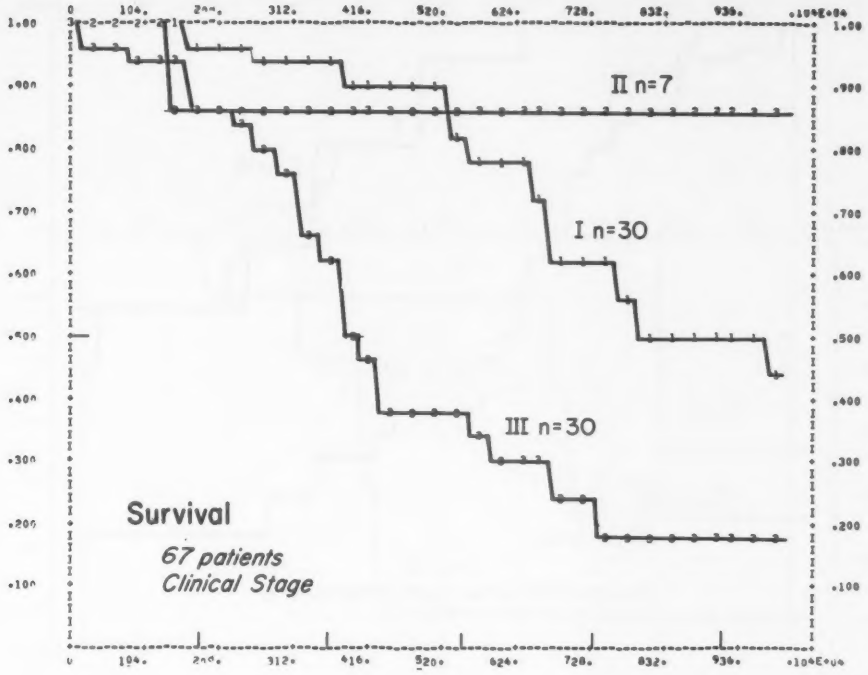



FIGURE 3.—Survival curves by clinical stage in 67 patients with testicular cancer (ordinate = accumulative proportions surviving; abscissa = days.

The decision was made early to use the Statistical Packing for the Social Sciences (SPSS-Versions 6.5 and 7.0) on the University CDC Cyber 74 computer for almost all aspects of data editing management and statistical analysis except the more complex multivariate analysis. This is the first time we have relied so heavily on SPSS in a large project. Using SPSS would be cumbersome, even with special FORTRAN routines, if we anticipated frequent or extensive data updating, but we plan only a single update. The computer center will soon acquire a data managing system (Scientific Information Retrieval System) which can interface with SPSS. This addition may prove helpful for the final update of the data, as well as other studies with frequent updates.

## PRELIMINARY EXAMINATION OF STAGING

One of the goals of analysis is the evaluation and improvement of the clinical staging of testicular cancer in order to make staging a more useful tool. The relationship between clinical stage and survival time was explored using actuarial life-table techniques and their associated survival curves to identify risk factors. In the process, many difficulties associated with partially retrospective data were encountered.

Clinical staging is a classification of the physical extent of the disease. At the time of orchiectomy, the newly diagnosed patient may be at any one of three stages, and subsequent treatment is selected partly according to stage (table 2). Survival curves of different groups of patients show the cumulative proportion of surviving patients in each group graphed against the time since diagnosis (figs 3–7). The figures are reproduced directly from the SPSS output with the addition of drawn lines and identification labels.

Figure 3 shows the survival curves for the entire group of patients in the preliminary file. The problem is obvious: patients with clinical stage I disease would be expected to survive longest, but at 2½ years (915 days) the cumulative proportion surviving is only 51%, whereas for stage II it is 86% and for stage III it is 18%. There are several possible explanations for this discrepancy.

*Statistical artifact associated with small numbers.*—There are only seven patients in the stage II group. Significance testing is not usually pertinent in a preliminary investigation, but in this case both a t test of the differ-

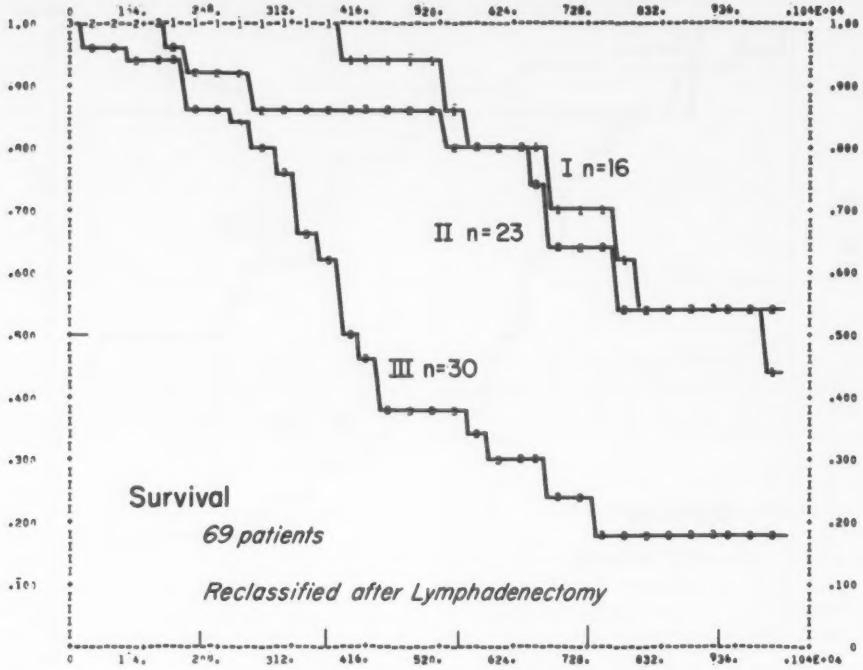



FIGURE 4.—Survival curves of 69 patients with stage reclassified after lymphadenectomy.

ence at 2½ years and Gehan's nonparametric test of differences in survival times (20) show significant differences at the 0.05 level. The problem appears to be more than that of numbers.

*Adequate staging may require biopsy of retroperitoneal lymph nodes.*—Twenty-five of the 30 patients initially diagnosed as having clinical stage I disease had a subsequent lymphadenectomy. Sixteen of these 25 patients (64%) were classified as having stage II disease after histologic study of the lymph nodes. In the survival curves by revised stage, the difference between stages I and II disappeared (fig 4). Stage I patients, however, still do not show the expected long survival.

*The patients at a referral hospital are not representative of the total patient population.*—Only 5% of the 78 patients had their orchiectomy at the University of Minnesota Hospital. Only 31% of patients with stage I disease and 44% with stage II disease had the lymphadenectomy at the University Hospitals. It seems that a stage I patient initially treated elsewhere and cured will not appear in our files. Recurrence in a stage I patient will lead to referral. The 20 patients who had either the orchiectomy or the lymphadenectomy at the University are more representative of the general population of patients with testicular cancer (fig 5). The numbers are small, but they show the expected ranking of survival by stage.

*Pathology, stage, and survival are confounded.*—Patients with seminoma as the primary tumor usually survive longer than those with other types of testicular cancer. In the current study the patients with nonseminomatous tumors have the expected pattern of survival (fig 6), whereas the eight stage I patients with seminoma (fig 7) have only 40% cumulative survival. It is clear that these eight patients do not represent a cross-section of all patients with stage I seminoma; they were referred to the University because they had recurrence of their cancer. Therefore, we are looking merely at the survival of patients with possible advanced disease who retrospectively were stage I at diagnosis. The patients with stage I seminoma who have done well tend not to be in our study at all, since they are never referred to the University Hospitals.

## COMMENT

The lesson is obvious: retrospective study data may contain peculiar biases, hence contradictory or surprising results should be fully explored, and conclusions should

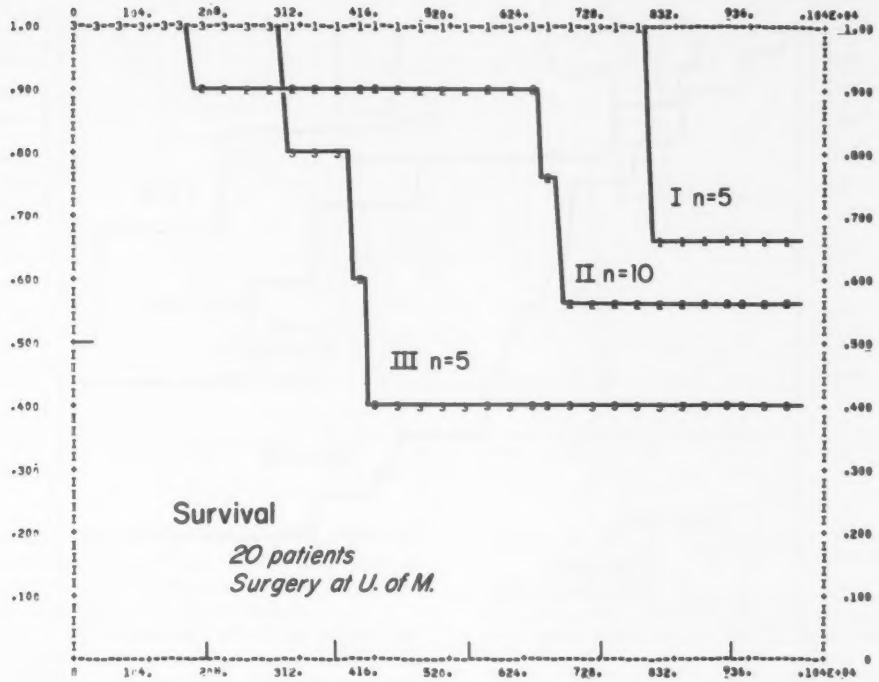



FIGURE 5.—Survival curves of 20 patients who had orchiectomy or lymphadenectomy at the University of Minnesota Hospitals.

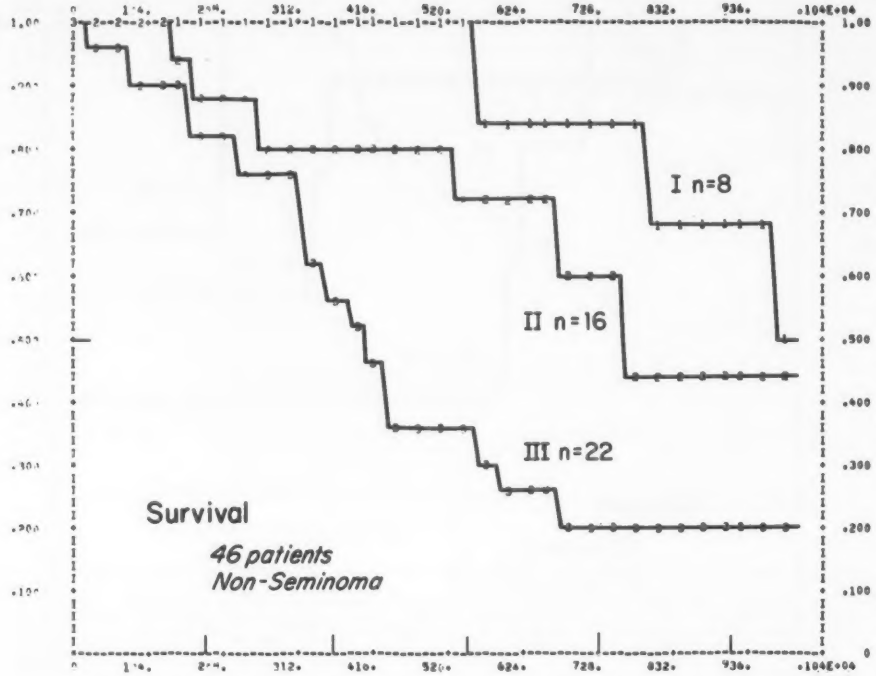
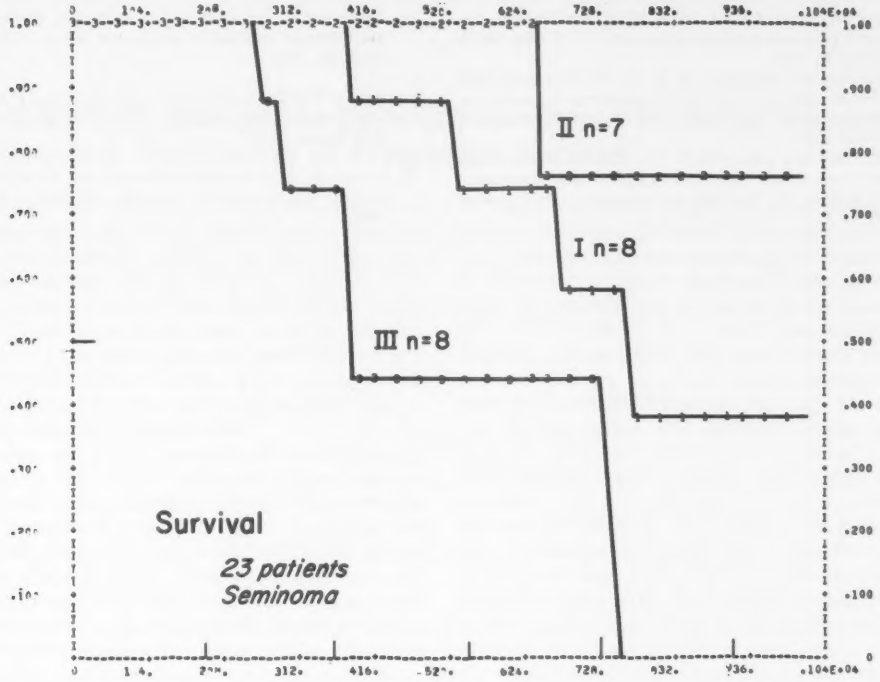



FIGURE 6.—Survival curves of 46 patients with nonseminomatous cancer of the testis.




be guarded until they can be supported by studies of appropriate design. Nevertheless, even though much of these data are prospective, analysis can make a valuable contribution to the design of future studies.

The small number of patients whom any center can enter into a clinical study of testicular cancer emphasizes the need for national pooling of clinical investigations and uniformity in the reporting of histology, staging, and results of treatment. Carcinoma of the testis is a major cancer that can be cured. To attain this goal we will need public education, professional education, assessment of risk factors, appropriate staging, and application of the most effective therapeutic regimens.

## REFERENCES


1. Twito DI, and Kennedy BJ. Treatment of testicular cancer. Ann Rev Med 26:235–243, 1975.
2. Brown JH, and Kennedy BJ. Mithramycin in the treatment of disseminated testicular neoplasms. N Engl J Med 272:111–118, 1965.
3. Curreri AR, and Ansfield FJ. Mithramycin-human toxicology and preliminary therapeutic investigation. Cancer Chemother Rep 8:18–22, 1960.
4. Yarbro JW, and Kennedy BJ. A comparison of the rate of recovery from inhibition of RNA synthesis in mouse liver and transplantable glioma. Cancer Res 27:1779–1782, 1967.
5. Kennedy BJ, Griffen WO, and Lober P. The specific effect of mithramycin on embryonal carcinoma of the testis. Cancer 18:1631–1636, 1965.
6. Samuels ML, Johnson DE, and Holoye PY. The treatment of stage 3 germinal cell neoplasia of the testis with bleomycin combination chemotherapy. Proc Am Assoc Cancer Res 14:23, 1973.
7. Samuels ML, Holoye PY, and Johnson DE. Bleomycin combination chemotherapy in the management of testicular neoplasia. Cancer 36:318–326, 1975.
8. Cvitkovic E, Cheng E, Whitmore WV, et at. Germ cell tumor chemotherapy update. Proc Am Assoc Cancer Res and ASCO 13:324, 1977.
9. Einhorn LH, and Donohue J. *Cis*-diamminedichloroplatinum, vinblastine, and bleomycin combination chemotherapy in disseminated testicular cancer. Ann Intern Med 87:293–298, 1977.
10. American Joint Committee for Cancer Staging and End Results Reporting. Manual for Staging of Cancer, 1977, p 127.
11. Bloomfield CD, Goldman A, Dick F, et al. Multivariate analysis of prognostic factors in the non-Hodgkin's malignant lymphomas. Cancer 33:870–879, 1974.
12. Kiang DT, Sundberg RD, Fortuny IE, et al. Multiple myeloma: clinical evolution. Minn Med 57:542–548, 1974.
13. Byar DP, Huse R, Bailar JC, et al. A multivariate model for predicting survival of prostatic cancer patients with time-censored data. JNCI 52:325, 1974.
14. Lange PH, and Fraley EE. Serum alpha-fetoproteins and human chorionic gonadotropin in the treatment of patients with testicular tumors. Urol Clin North Am 4:393–406, 1977.

15. Nochomovitch LE, De La Torre RFE, and Rosai J. Pathology of germ cell tumors of the testis. Urol Clin North Am 4:359–378, 1977.
16. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 34:585–612, 1976.
17. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer 35:1–39, 1977.
18. Armitage P, and Gehan EA. Statistical methods for the identification and use of prognostic factors. Int J Cancer 13:16–36, 1974.
19. Bishop YMM, Fienberg SE, and Holland PW. Discreet Multivariate Analysis: Theory and Practice. Boston, MA, MIT Press, 1975.
20. Gehan E. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52:203–224, 1965.

# DISCUSSION III: Designs to Optimize Combined-Modality Therapy and Selection and Verification of Prognostic Factors*

**Dr. Breslow:** Drs. Rubin and Salazar have presented a comprehensive review of the introduction of radiation therapy into cooperative group programs and the logic of combined-modality therapy. When they present their final tally of the number of studies using radiation, I hope they will include those which contain a standard radiation protocol even if the protocol does not specifically address some aspect of radiation therapy. For example, one of their tables did not list three of our studies which do have a standard radiation therapy protocol.

The authors put a welcome emphasis on documenting and analyzing the sites of treatment failure, whether these be local, nodal, or distant. As part of the combined modality approach it is most important to identify the ability of different treatments to control different aspects of disease. There is an increasing recognition that such prognostic factors as histologic type and regional lymph node involvement may signal specific types of failure. Such information is potentially very useful as a means of optimizing the treatments selected for each patient. However, in this context I would like to ask one general question: what is the logic of reserving the most intensive therapy for high-risk patients who may have the least potential to benefit from it?

Both Dr. Weiss and, earlier, Dr. Byar have questioned the clinical relevance of current models for immunotherapy involving experimentally induced tumors in animal species. I too have wondered at the proliferation of unimaginative protocols using such nonspecific therapies as MER and BCG for the treatment of cancers occurring in a wide range of sites and stages. If the scientific community considered the animal results more carefully, as Dr. Weiss suggests, we should see a reversal of this trend and the development of immunotherapy protocols which have a firmer scientific foundation. Otherwise, I fear that the lack of progress made with the present "shotgun" approach will bring strong pressure for the termination of this line of investigation.

Regarding Dr. Weiss' recommended significance level of $P = 0.05$ for smaller trials, it might be well to use $P = 0.01$ or even less as a guard against the tendency to single out and report only the significant results from several similar small trials.

Drs. Goldman and Kennedy describe a search for prognostic factors in testicular cancer and identify many of the issues which affect prognostic studies generally. They observe that testicular cancer is the most common neoplasm in white males between the ages of 20 and 34 years. In fact there are usually two peaks in the age-incidence curve for testicular cancer, one at age 30 years and the other later in life. In the United States and the United Kingdom, the mortality rates among young adults with testicular tumors have been increasing, whereas such rates for the elderly have been declining. The disease is rare among African and American blacks, and it is strongly associated with several congenital disorders. These features seem to indicate that a genetic or developmental event puts some males at high risk from early life, but thereafter the appearance of the tumor is possibly mediated by environmental factors. The reason for pointing out these facts is to emphasize that the astute clinician is often in the best position to recognize clues to the origin of the disease. Although his first mission must be the eradication of the disease with treatment, the oncologist can make an additional contribution when he helps to identify the causes and thus assists in prevention. An increased awareness of such issues could eventually relieve some of the frustrations which arise when we try to conduct case-control studies based on patients' medical history and find that much of the relevant information has not been recorded.

I was struck by the observation that mithramycin was more effective for certain cell types than others. The identification of such "treatment-covariate interactions" is extremely important as one tries to identify optimal treatment strategies for individual patients. However, this general area also raises serious methodologic issues, particularly that of multiple comparisons. In treatment-oriented studies it it tempting to divide the patients into subgroups on the basis of histology, extent of disease, and other factors. Since at least some of the variation in treatment effectiveness from subgroup to subgroup is due to chance, we must be cautious when claiming that a particular treatment is effective for one part of the patient population and not another. The most convincing evidence for such an effect lies in the replication of the observation in several studies.

I was impressed by the fact that > 50% of the patients with advanced testicular disease could be cured by modern therapy. In children and young adults it is relatively easy to document such cures, since there are few competing causes of death. The percentage cured may be estimated from the tail of the survival curve, which should remain virtually flat after 2, 3, or 5 years for those

*Reprint requests to: Dr. Norman Breslow, Department of Biostatistics, University of Washington, Seattle, WA 98105.

patients in whom the disease has been completely eradicated. One hopes that the curves shown by the authors will, in fact, show such flattening when the time scale is extended. Otherwise it would be premature to talk of "cure" here since the curves shown only went out to about 3½ years. It is much harder to identify the treatment successes among elderly patients who are subject to many causes of death.

Whereas many of the classical prognostic variables, including histologic type, age, and stage of disease at diagnosis, are fixed for each patient, some of the newer biologic markers may vary substantially over time. This variation represents both an opportunity and a challenge to the statistician, since it allows the possibility of altering the patient's prognosis as additional information becomes available during the followup period. This is one place where Professor Cox's life-table regression model with time-dependent covariates can be profitably employed. It permits one to model rates of recurrence or mortality as functions of the patient's accumulating history.

In the written version of their paper, Drs. Goldman and Kennedy referred to the cross-checking of contradictory data. This is a problem we all have to deal with. It is not unusual to find a discrepancy between the lymph node evaluations of surgeon and pathologist, for example. In some instances, this may be due to an error in the labeling and handling of the surgical material. More often it reflects real differences in the outcome of gross versus microscopic study of the same tissue. In such cases, one tends to let the pathologist have the last word. However, it has been interesting to note in our own data on Wilms' tumor that patients with lymph nodes declared positive by the surgeon yet negative on microscopic study seemed to have a prognosis as bad as that for patients in whom the nodal metastases were verified microscopically.

When one is dealing with a curable disease, such as testicular cancer, I have some doubts as to whether we should employ the conventional life-table methodology and related multivariate statistical procedures. These methods may identify progonostic factors by their effects on early variations in the shape of the survival curve, rather than on the ultimate fraction cured. In such cases it makes more sense simply to treat the patient's survival status at 2, 3, or 5 years as a binary response indicator which is related to the prognostic factors using, for example, linear logistic regression. Of course, in this approach one must have followed all patients for the minimum risk period. The often observed discrepancies between the results of interim and final analyses may reflect an early separation of the survival curves which later come together.

I agree that adequate staging may require biopsy of the lymph nodes. Generally speaking, valid statistical procedures require that each of the data items be uniformly defined and measured for all patients. This implies that the extent of surgical exploration and the utilization of other diagnostic procedures will be standardized. However, because the degree of exploration is as often linked to clinical impressions as to the extent of disease, such a requirement may well conflict with standard practice and could even threaten the patient's well-being. Thus it may be necessary to make separate evaluations of prognosis, depending on which set of diagnostic procedures was actually employed.

The authors themselves have identified the most serious problem with their data, the fact that some patients were entered into the study only upon referral for treatment of recurrent or metastatic disease. Attempts to include such patients in the study of prognostic factors operative at the time of diagnosis would introduce severe bias. For this purpose it would be better to utilize only the newly diagnosed and treated patients. However, others might be incorporated into a study of factors that predict for survival among patients with late recurrent disease.

# Designing for Cancer Clinical Trials: Selection of Prognostic Factors[1]

**Byron Wm. Brown, Jr**[2,*]

## SUMMARY

**This paper reviews the pros and cons for stratifying on a number of variables when randomizing patients to treatments in cancer clinical trials. Arguments in favor of randomization focus on the increased precision achieved. Arguments against stratification focus on the complexity of randomization procedures and the fact that post hoc statistical adjustment can achieve nearly the same precision as stratification. Although arguments on both sides have merit, newer methods of adaptive randomization would seem to shift the balance toward the use of predictive factors in achieving balance among treatment groups. In the Northern California Oncology Group, the Efron-biased coin method of randomization is being used to balance treatment groups on as many as four to six prognostic variables. Treatment assignment is made by telephone to the Statistical Center, where the assignment is determined by computer, taking into account previous assignments and the prognostic characteristics of the patient to be assigned.**

[Cancer Treat Rep 64:499–502, 1980]

Prognostic factors have three important uses in the design and analysis of clinical trials. First, prognostic factors are used to define the eligibility of the patient for the trial. Second, these factors are used to group patients into homogeneous strata or blocks, random assignment of treatments then being carried out within each stratum, to assure balance in the numbers on each treatment arm within each stratum. Third, these factors are used in the statistical analysis of the data, to assure that comparisons of results in different treatment arms are adjusted to account for variations in prognosis at the time of random assignment.

It should be noted that, even if a prognostic factor is not used in deciding the eligibility or treatment allocation, but is used in the analysis of the data at the end of the study, the prognostic factor must be specified at the start of the study so that it can be carefully defined in the protocol and reliably and accurately observed and recorded at baseline for each patient.

Though the use of prognostic factors to define the eligibility of patients for a clinical trial is an important topic, my discussion will be confined to the other two uses of prognostic factors, in random assignment and in analysis. There is a controversy now, especially among statisticians, as to whether prognostic factors should be used in random assignment of treatments to patients. One camp advocates heavy use of prognostic factors to assure that the patients assigned to the two arms show close balance in prognoses at baseline. The other camp maintains that the random assignment should be made without regard to prognostic factors, and that fair comparison of treatment effects can be achieved through statistical adjustment of the results. I would like to review the arguments that have been given on either side, and then give my own view.

It is clear that if prognostic factors are used in the random assignment of treatments to patients, the treatment groups will be more comparable than if the prognostic factors had been ignored. Many statisticians and clinicians argue that it is essential to have close comparability of the treatment groups with regard to prognostic factors. For example, Staquet (1) says that "groups of patients to be compared after treatment must be alike"; otherwise, he says that erroneous conclusions can result,

[1]Supported by Public Health Service grants CA-21744-01 (National Cancer Institute) and GM-21215 (National Institute of General Medical Sciences), National Institutes of Health, Department of Health, Education, and Welfare.

[2]Division of Biostatistics, Stanford University School of Medicine, Stanford University Medical Center, California.

**Reprint requests to:* Byron Wm. Brown, Jr, PhD, Division of Biostatistics, Department of Family, Community and Preventive Medicine, Stanford University Medical Center, Stanford, CA 94305.

due to "subtle bias." Zelen (2) says that we should "plan experiments so that the process of randomization takes into account the major prognostic variables so that they are equally allocated among the treatment groups." Failure to do this "makes the comparisons between treatment groups relatively imprecise." Pocock and Simon (3) say that if one treatment group, through random assignment, chances to have more patients with poor prognostic factors, "considerable bias could result in any comparison of treatments." They go on to say that "it may be possible to remove such bias by introducing the (prognostic factor) as a covariate in the analysis," but they note some difficulties in such post hoc statistical adjustment. Armitage and Gehan (4) are somewhat ambiguous in their advice on the use of prognostic factors in randomization, but do say that, in a small to moderate clinical trial (ie, one with $\leq$ 100 patients) the patients in the two treatment groups "could be non-comparable and a meaningful comparison would depend on having a valid quantitative model relating response to prognostic factors." They conclude that "knowledge of prognostic factors is important in planning small to moderate-sized randomized control studies. . .," and the implication is that prognostic factors must be used in the randomization procedure. However, they do call for more study of the question.

The statisticians who profess by word or action the necessity of using prognostic factors to balance the treatment groups use a variety of terms in arguing the case. Their choice of words suggests that there may be some disagreement among them as to whether the problem of non-comparability is one of bias (that is, invalidity or meaninglessness of the results), or simply a question of less power or sensitivity in a study that ignores prognostic factors in assigning treatments. Among those statisticians whom I have quoted, Staquet, Pocock and Simon, and Armitage and Gehan seem to hold to the belief that the results might indeed be invalid or meaningless if prognostic factors are not used in assigning treatments, whereas Zelen states that the problem is one of insensitivity or diminution of power.

There is another thread running through the arguments for balancing treatments on prognostic factors, and that has to do with the possibility that the difference in efficacy between the two treatments under study may vary from one prognostic group to the next. For example, treatment A may only be more effective than treatment B for patients with cancer of one histologic type. Thus, many statisticians advocate testing and estimation of stratum by treatment interactions. It is obvious that estimates of differences in treatment efficacy for patient subgroups will be less precise than estimates of the overall difference based on all patients, and balance in numbers and prognostic variables for the subgroups could greatly benefit such analyses. Zelen, and also Armitage and Gehan, have argued for balancing on this ground.

Now let us turn to the arguments against the use of prognostic factors in randomization. For many of us who consult on the design of clinical trials, the keyword is simplicity, and we believe that a randomized clinical trial that uses a number of stratification variables in the randomization procedure is not worth the trouble and the risk of mismanagement that such complicated procedures entail. Recently a group of ten British and American statisticians, led by Peto et al (5,6) set down the arguments for this view in an elegant discussion of the design and analysis of clinical trials in the *British Journal of Cancer*. It is especially significant that their arguments were directed to the design of clinical studies of cancer, since clinical scientists specializing in the study of cancer are more inclined than others to the heavy use of prognostic factors in randomization procedures.

The following are arguments these ten statisticians marshal against the use of prognostic factors. First, even if prognostic factors are ignored in making the treatment assignments, randomization will tend to balance the treatment assignments within each of the stratification cells. In fact, it is easily shown that if there are N patients in a given cell, the amount of information achieved by complete balance, ie, exactly one half of the N patients assigned to each treatment, can be achieved on the average by randomly assigning N + 1 patients to the treatment arms, by the flip of a fair coin, without regard to balance.

The second argument of Peto's group concerns the complexity of the balancing operation. They argue that such complexity may actually discourage patient accrual in a multicenter trial and result in slower progress or even failure of the trial altogether, for lack of physician cooperation.

In their third argument, they point out that the use of prognostic factors to balance the treatment groups does not obviate the necessity for using these variables in the statistical analysis of the results. In fact, they claim that just the opposite is true. If a variable is used in the randomization procedure to balance the treatment groups overall, and then is ignored in the analysis, the level of significance ($P$ value) computed will not be valid. It will be larger than the $P$ value calculated on the basis of the method of randomization actually used, so the sensitivity of the study will be adversely affected. Thus, if a randomization procedure is used that is based on a multitude of prognostic factors, these factors must be used in the analysis, or the advantage of the more complex procedure is lost in the evaluation of the results.

My own view of the preceding arguments is as follows: first, I find the arguments against the use of prognostic factors in randomization of treatment assignments logically valid but not persuasive. Why settle for a design that will do well, on the average, when one can purposely balance and make sure of doing well in the study at hand? Why settle for less information per cell, even if the information sacrificed per cell amounts only to the equiv-

alent of one patient per cell? In fact, one patient per cell, in a highly stratified experiment, may amount to a substantial number of patients. The arguments concerning the complexities of using more intricate randomization and, later, the more complex statistical analyses which will be required, do not put me off either. The situation may well be quite different for others, and complicated randomization schemes may be out of the question for some investigative groups, but in the Northern California Oncology Group, registration and randomization of each patient are phoned to the Statistical Center, where the name of the patient and his prognostic factors and institution are entered in the computer while the data manager in the Center is talking with the physician or clinical assistant; the computer then responds with the treatment assignment. The whole procedure takes < 5 minutes. The assignment is randomly selected by the computer, using a random number generator and a forced balancing routine adapted from procedures described by Efron (7) and resembling the adaptive procedures described by others like Pocock and Simon (3) and Wei (8). The result is a set of assignments that will balance the treatment groups with regard to each of the prognostic factors and within any stratum cell that has more than a very few patients. As to the complexity of analyses, Peto and his colleagues have shown us a variety of simple and efficient ways of incorporating a large number of factors into our analyses of survival data; the problems in statistical and computational theory and practice are routinely handled now by most statistical centers, and I believe that we need argue no longer that the more appealing balanced design should not be used because analyses reflecting the design are difficult or unavailable. It is even clear that appropriate permutation tests matching the adaptive designs are achievable with the use of a large computer, though Efron (9) has argued that we need not go to such exotic lengths.

Thus, I would advocate the use of prognostic factors in allocation of treatments to patients. I would like to disclaim, however, certain of the arguments that have been advanced for this view. Contrary to the views of Staquet, Armitage, Gehan, and Pocock and Simon, the completely randomized experiment is not an invalid experiment, even if the treatment groups do not prove comparable on all prognostic factors after all assignments are made. In fact, the significance test $P$ value, computed *without* regard to the prognostic factors, or, better, computed *with* adjustment for (ie, conditional on) the prognostic factors, is directed to the very question of whether the difference between results for the two treatments can be attributed solely to the non-comparabilities that can come about through random assignment, or whether the treatment difference observed must be attributed in part to a real difference between the efficacies of the two treatments. The purpose of the randomization is to make the calculation of the $P$ value valid, and the purpose of the $P$ value is to assess the evidence for a real treatment difference in the face of groups that are *necessarily* different with regard to prognostic factors. Thus, the purpose of balancing is not to provide a less biased or more meaningful comparison, but to provide a more precise comparison, as Zelen argues. This added precision, however slight, comes about through the smaller standard errors that go along with equal allocations of patients, cell by cell, to the several treatments.

A word about the estimation of interactions is in order here. Although it is appealing to talk of the advantages of balance for the estimation of treatment differences for individual cells or strata, it is not important for most studies. Most cancer studies are designed to yield marginally precise comparisons of treatments when averages are taken over all strata, and there is little hope of obtaining estimates with satisfactorily small standard errors for individual cells or strata, even when treatment arms are in perfect balance. Thus, although I favor balancing, I do not feel that estimation of treatment differences for individual strata is a strong argument for doing so.

My feeling, then, is that there is some slight advantage in precision, ie, smaller standard errors, when prognostic factors are taken into account in the randomization procedure. And there is little disadvantage. Furthermore, results of an experiment in which the treatment groups show balance, with regard to each prognostic factor and in each cell with more than a few patients, carry an element of persuasiveness in themselves, especially for the clinician (and statistician) who is not persuaded of the validity of statistical adjustment for imbalances.

There is not the time here to discuss the choice of prognostic factors or the methods of randomization that should be used in order to balance the treatment groups with regard to these factors. But the two considerations go together. Since one of the reasons for balancing is to secure a balance that will reassure less statistically oriented auditors, it is important to achieve balance with regard to the well-accepted prognostic factors for the disease under study, even if there are half a dozen or more, and even though it is almost certain that the marginal gain in precision in adding the last factor is negligible. Now randomization within each cell, when there are half a dozen stratification factors, is hopeless, since many cells will have only one patient. Balance on each prognostic factor, however, can be achieved by some sort of adaptive scheme, where each assignment is made with probability favoring the treatment which will bring the various factors back in balance between the two arms. I have mentioned a number of such adaptive schemes, and I favor this approach.

In summary, then, I believe that in designing a clinical trial the choice of prognostic factors should be made on the basis of the well-accepted dogma of the disease under study; that one should not hesitate to use as many as half

a dozen such factors; and that these factors should be used explicitly in making treatment assignments, following some sort of adaptive or forced balancing procedure. The result will be a study with slightly greater precision and credibility than a study in which the treatments are randomly assigned without regard to prognostic factors.

## REFERENCES


1. STAQUET MJ, ed. Preface. *In* Cancer Therapy: Prognostic Factors and Criteria of Response. New York, NY, Raven Press, 1975.
2. ZELEN M. Importance of prognostic factors in planning therapeutic trials. *In* Cancer Therapy: Prognostic Factors and Criteria of Response (Staquet MJ, ed). New York, NY, Raven Press, 1975.
3. POCOCK SJ, and SIMON R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 31:103–115, 1975.
4. ARMITAGE P, and GEHAN EA. Statistical methods for the identification and use of prognostic factors. Int J Cancer 13:16–36, 1974.
5. PETO R, PIKE MC, ARMITAGE P, ET AL. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 34:585–612, 1976.
6. PETO R, PIKE MC, ARMITAGE P, ET AL. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer 35:1–39, 1977.
7. EFRON B. Randomizing and balancing a complicated sequential experiment. *In* Biostatistics Case Book. Technical Report No 21. Stanford, CA, Division of Biostatistics, Stanford University, 1976, vol 1, pp 21–31.
8. WEI LJ. The adaptive biased coin design for sequential experiments. Ann Stat 6:92–100, 1978.
9. EFRON B. Forcing a sequential experiment to be balanced. Biometrika 58:403–417, 1971.

# Designing for Nonparametric Bayesian Survival Analysis Using Historical Controls [1]

John Van Ryzin [2,*]

## SUMMARY

**This paper gives a method for choosing the number of patients $N_0$ out of N available patients to be randomized to current controls in a two-arm study when comparison of nonparametric survival curves is the anticipated method of data analysis. The criterion imposed is that of choosing $N_0$ to minimize the posterior variance of the difference between the current control and experimental survival curves. A nonparametric Bayesian argument incorporating the survival curve of available historical controls establishes the criterion. Formulas and tables which facilitate this computation are presented.**

[Cancer Treat Rep 64:503–506, 1980]

When designing a clinical trial to compare a treatment group with a control group, the first problem to be faced is the choice of an appropriate control group. Within the same organization or within the same cooperative group, it has been recognized that besides randomized current controls it may well be beneficial to also use historical controls. This point has been discussed in some detail by Gehan and Freireich (1), Meier (2), and Pocock (3,4), to mention but a few. When the analysis is to compare the treatment group with the control group by using nonparametric survival curves, this paper describes a method of designing the sample size for the control group relative to the treatment group, based on combining the control group survival curve with that of a historical control survival curve available from a previous study. This method and its use is given in the sections, "Derivation of Design Using Historical Controls" and "Use and Interpretation of Design." The following section discusses methods of nonparametric survival analysis.

## NONPARAMETRIC SURVIVAL ANALYSIS

### Kaplan-Meier Curves

A commonly used method of nonparametric estimation of survival curves is that of Kaplan and Meier (5). Their method is briefly described. Let N patients be entered in a study under a given treatment. Suppose $X_1', \ldots, X_N'$ are the true lifetimes (or disease-free interval times, or times to remission) of the N patients and $X_1^* \ldots, X_N^*$ are the N followup times of the patients under study (usually, $X_i^*$ is the time from entry of patient i to close of the study). At the end of the study, the investigator observes

$$X_i = \min(X_i', X_i^*) \qquad [A]$$

and

$$\delta_i = \begin{cases} 1 & \text{if } X_i' \leq X_i^* \\ 0 & \text{if } X_i' > X_i^* \end{cases} \qquad [B]$$

If $\delta_i = 0$, then $X_i (= X_i^*)$ is said to be a "censored lifetime" and if $\delta_i = 1$, then $X_i (= X_i')$ is a "true lifetime."

Using these data we wish to estimate the survival curve

$$S(t) = \Pr\{X_i' > t\}, \qquad [C]$$

under the assumption that $X_1', \ldots, X_N'$ are a random sample from the underlying true survival curve S(t). The Kaplan-Meier estimator of the survival curve S(t) can be written as

$$\hat{S}(t) = \frac{N^+(t)}{N} \prod_{\{i:X_i \leq t\}} \left\{ \frac{N^+(X_i)+1}{N^+(X_i)} \right\}^{1-\delta_i} \qquad [D]$$

where the product is over all $X_i \leq t$ and $N^+(x)$ = number

[1] Supported by Public Health Service grant GM-23129 from the National Institute of General Medical Sciences, National Institutes of Health, Department of Health, Education, and Welfare.

[2] Division of Biostatistics, Columbia University, New York, NY.

[3] *Reprint requests to:* Dr. John Van Ryzin, Division of Biostatistics, School of Public Health, Columbia University, 600 W 168th St, New York, NY 10032.

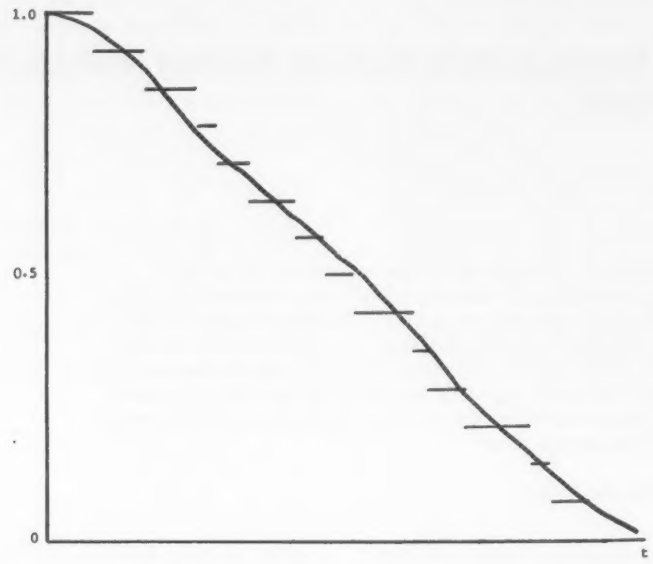



FIGURE 1.—A typical survival curve S(t) and its Kaplan-Meier estimate $\hat{S}(t)$.

of $X_i > x$ and $\delta_n$ is taken as 1. Equation [D] is not the usual form of writing the Kaplan-Meier curve, but it is exactly the same as the usual form (Susarla and Van Ryzin [6]).

Figure 1 illustrates a typical survival curve S(t) and its estimator $\hat{S}(t)$.

## Bayesian Survival Curves

Suppose now that we have historical information about survivals of the patients which can be represented by a survival curve $S_H(t) = \beta(t)/\beta$, where $\beta(t)$ is a nonnegative, nonincreasing function of t with $\beta = \beta(0) > 0$ and $\beta(\infty) = 0$. For example, if we have a set of historical control data on $N_H$ patients we may wish to take $S_H(t) = \hat{S}_H(t)$, where $\hat{S}_H(t)$ is the Kaplan-Meier curve [D] based on the historical controls of the type [A] and [B] for the $N_H$ patients.

Using such prior survival curve information contained in the parameter curve $\beta(t)$, Susarla and Van Ryzin (6) have used the nonparametric Bayesian methods of Ferguson (7) to derive a nonparametric Bayesian survival curve estimator based on the data [A] and [B] and the prior survival curve $\beta(t)/\beta$. Their estimator if there are no ties in the data becomes

$$\hat{S}_\beta(t) = \frac{\beta(t) + N^+(t)}{\beta + N} \prod_{\{i:X_i \le t\}} \left\{ \frac{\beta(X_i^-) + N^+(X_i) + 1}{\beta(X_i^-) + N^+(X_i)} \right\}^{1-\delta_i} \qquad [E]$$

where the product is over all $X_i \le t$, $\beta(s^-) = \lim$ of $\beta(t)$ as t increases to s, $N^+(x)$ is as in [D] and $X_i$ and $\delta_i$ as in [A] and [B].

The Bayesian nonparametric estimator $\hat{S}_\beta(t)$ allows combining historical data through $\hat{S}_H(t) = \beta(t)/\beta$ with current data given by [A] and [B].

It is interesting to examine $\hat{S}_\beta(t)$ in the case of no censoring, that is, when all lifetimes $X_i$ are available. In that case,

$$\hat{S}_\beta(t) = \frac{\beta}{\beta+N} S_H(t) + \frac{N}{\beta+N} \hat{S}(t) \qquad [F]$$

where $S_H(t)$ is the historical survival curve and $\hat{S}(t)$ is the Kaplan-Meier estimated curve for the data of [A] and [B].

From equation [F] we have the interpretation that $\beta/N$ is the ratio of weight (or belief) in the historical survival

curve relative to the sample size N of the current data. If the historical and current data carry equal weight, $\beta = N$, then $\hat{S}_g(t) = \frac{1}{2}(S_H(t) + \hat{S}(t))$ in [F], an average of the historical and current survival curve.

We now introduce the design problem for comparing a control and treatment survival curve when an historical control survival curve is available.

## DERIVATION OF DESIGN USING HISTORICAL CONTROLS

Let $S_0(t)$ and $S_1(t)$ be the true survival curves for the control and treatment groups in a planned clinical trial utilizing N patients. Suppose $N_0$ and $N_1 = N - N_0$ patients are to be randomly assigned to the control and treatment groups respectively. In designing such a study we shall assume all patients will be followed until death. Suppose that prior survival curves are given for both the treatment and control groups by $\beta(t)/\beta$ and $\alpha(t)/\alpha$ respectively. If one assumes $1-S_0(t)$ and $1-S_1(t)$ have independent Dirichlet processes with parameters $\beta(t)$ and $\alpha(t)$, and one uses the theory developed by Ferguson (7), the survival curves $1-S_0(t)$ given the data and $1-S_1(t)$ given the data are independent posterior Dirichlet processes with parameters $\beta(t)+N_0-N_0^+(t)$ and $\alpha(t)+N_1-N_1^+(t)$, where $N_i^+(t)$ = number of lifetimes $> t$ among the $N_i$ patients, $i = 0,1$.

To design the experiment we must choose $N_0$, the number randomly assigned to the control group. In order to do this we shall minimize the posterior variance (Post Var) of $S_0(t)-S_1(t)$ based on the above theory. Using properties of the Dirichlet distribution, it can be shown that

$$\text{Post Var}(S_0(t)-S_1(t)) = \frac{(\beta(t)+N_0^+(t))(\beta+N_0-\beta(t)-N_0^+(t))}{(\beta+N_0)^2\,(\beta+N_0+1)} + \frac{(\alpha(t)+N_1^+(t))\,(\alpha+N_1-\alpha(t)-N_1^+(t))}{(\alpha+N_1)^2\,(\alpha+N_1+1)} \cdot \quad [G]$$

Now usually historical survival information is not available for the treatment group. Therefore, we minimize this posterior variance with respect to $N_0$ under the assumption that we have *no prior* survival curve for the experimental group; that is, we take $\alpha \longrightarrow 0$ in [G] (where $\alpha(t) \leq \alpha$ for all t). However, assuming historical controls we take $\beta(t) = \beta\hat{S}_H(t)$, where $\hat{S}_H(t)$ is the Kaplan-Meier survival curve based on $N_H$ historical control patients. With these substitutions, the resulting posterior variance after some algebraic simplification becomes

$$\text{Post Var}\ (S_0(t) - S_1(t)) = \left(\frac{\beta}{\beta+N_0}\,\hat{S}_H(t) + \frac{N_0}{\beta+N_0}\,\hat{S}_0(t)\right)\left(\frac{\beta}{\beta+N_0}\,(1-\hat{S}_H(t)) + \frac{N_0}{\beta+N_0}\,(1-\hat{S}_0(t))\right) \cdot \frac{1}{\beta+N_0+1} + \frac{\hat{S}_1(t)(1-\hat{S}_1(t))}{N-N_0+1} \cdot \quad [H]$$

where $\hat{S}_i(t) = N_i^+(t)/N_i$, $i = 0,1$ are the Kaplan-Meier curves with no censoring for the current control and experimental groups respectively.

Since we cannot minimize [H] simultaneously for all t, we choose for design purposes to minimize [H] for the median survival time of the historical control group, that is, at $t = t_H$ such that $\hat{S}_H(t_H) = \frac{1}{2}$. Also, as we saw in discussing equation [F], $r = \beta/N_0$ represents the ratio of belief in the historical and current data controls. Our goal then is to minimize [H] with $t = t_H$ and $\beta = rN_0$, which, after algebraic simplification is

$$f(N_0) = \text{Post Var}\left(S_0(t_H)-S_1(t_H)\right) = \frac{a_1^2}{N-N_0+1} + \frac{a_0^2 + \frac{r(r+2)}{4}}{[(r+1)N_0+1]\,(r+1)} \quad [I]$$

where $a_i^2 = \hat{S}_i(t_H)(1-\hat{S}_i(t_H))$, $i = 0,1$. If one minimizes this by setting the derivative $f'(N_0) = 0$, one obtains

$$N_0 = N_0(r) = \frac{N}{1+(1+r)h(r)} + \frac{1-h(r)}{1+(1+r)h(r)} \quad [J]$$

where

$$h(r) = \sqrt{\frac{4a_1^2(1+r)}{r(2+r)+4a_0^2}}\,. \quad [K]$$

Formulas [J] and [K] give us the number $N_0$ of the N patients to randomly assign to current controls.

## USE AND INTERPRETATION OF DESIGN

The formula [J] which chooses $N_0$ to minimize the posterior variance of the difference of the two curves $S_0(t)$ and $S_1(t)$ looks complicated to use. It depends on the quantities $a_i^2$, $i = 0,1$ which are calculated from patient data not available at design time. However, since our interest is in comparing the two survival curves, we design the study under the assumption that $S_0(t) = S_1(t)$. Under the assumption $\hat{S}_0(t_H)(1-\hat{S}_0(t_H)) = \hat{S}_1(t_H)(1-\hat{S}_1(t_H))$ by large sample properties of the survival curve estimators (*see* Peterson [8]).

Let $\hat{S}_0(t_H)(1-\hat{S}_0(t_H)) = \frac{1}{4} - b^2$, $0 \leq b \leq \frac{1}{2}$, where b represents the bias of the historical controls from the randomized controls at $t = t_H$, that is, $b = |\hat{S}_H(t_H)-\hat{S}_0(t_H)|$. With these simplifications the formula for the optimal choice of $N_0$ is given by [J] with

$$h(r) = \sqrt{\frac{(1-4b^2)(1+r)}{r(2+r)+(1-4b^2)}}\,. \quad [L]$$

We now examine the solution for $N_0$ given by equations [J] and [L].

*Case 1: Using No Historical Controls* ($r = 0$)

If we take $r = 0$, $\beta = rN_0 = 0$, then the estimate $\hat{S}_0(t) = \hat{S}_0(t)$ (*see* [F]) and no historical control survival curve is used. Note that in this case, $h(0) = 1$, and the optimal

choice of $N_0$ by formula [J] is to take $N_0 = N/2$. Hence, if *one uses no historical control survival curve a priori, the optimal number of randomized controls is one half the number of available patients under study.*

*Case 2: Increasing Reliance on Historical Controls (r approaches ∞)*

As r approaches ∞, h(r) approaches 0 in [K] and [L] and $\sqrt{1+r}$ h(r) approaches the limit $\sqrt{(1-4b^2)/(2-4b^2)}$. Therefore, from formula [J], we have $N_0(r)$ approaches 0. Thus, since $r = \beta/N_0$ represents the relative ratio of the weight put on historical controls versus current controls, we have that as r increases to ∞, the optimal number of current controls decreases to 0. Thus, $N_0 = 0$ is equivalent to taking $\hat{S}_0(t) = \hat{S}_H(t)$ and we have *complete reliance on historical controls.*

*Case 3: Intermediate Values of r*

Tables 1–4 give the optimal solution for $N_0$, the number of randomized controls, based on the values of $r = \beta/N_0$ for r = 1, 3, 5, and 10. The values of b = 0, 0.1, 0.2, 0.3, and 0.4 and N = 100, 300, 500, and 1000 are included in each table.

As an example of using these tables, suppose we wish to give equal weight to the historical and current controls and we take $r = 1$ or $\beta = N_0$. Then, if we suspect a small bias of less than, say, 20% of the historical value $0.5 = \hat{S}_H(t_H)$, we would use the column b = 0.1(0.20 × 0.5) of table 1 yielding, say, 126 controls and 174 treatment patients for a study with N = 300 patients.

## CONCLUSION

Even though the *design criterion* in this paper is based on assuming all current patients (control and treatment groups) are followed until death, this is *not* a requirement to carry out the analysis. More complicated design criteria based on the posterior variance of a mixture of Dirichlet processes could be developed (*see* Blum and Susarla [9]).

The choice of r when using formula [L] should probably depend on $N_H$, the size of the historical control patients. The ratio r should be larger when more historical control data are available. This dependence, together with the tradeoff for the bias b, seems worth further investigation.

Since analysis of survival data using the Kaplan-Meier curves is often the method of analysis chosen in randomized clinical trials, it is hoped that this paper will encourage the use of historical controls in ongoing cooperative group studies wherein the bias from historical controls may be negligible. As seen from tables 1–4, considerable savings in the number of needed current controls will be available in such a setting.

TABLE 1.—$N_0$ for r = 1

| N\b | 0 | 0.1 | 0.2 | 0.3 | 0.4 |
|---|---|---|---|---|---|
| 100 | 42 | 42 | 43 | 46 | 52 |
| 300 | 124 | 126 | 129 | 137 | 154 |
| 500 | 207 | 209 | 215 | 229 | 257 |
| 1000 | 414 | 418 | 431 | 458 | 513 |

TABLE 2.—$N_0$ for r = 3

| N\b | 0 | 0.1 | 0.2 | 0.3 | 0.4 |
|---|---|---|---|---|---|
| 100 | 33 | 34 | 35 | 44 | 45 |
| 300 | 100 | 102 | 106 | 133 | 135 |
| 500 | 167 | 169 | 176 | 221 | 225 |
| 1000 | 333 | 338 | 352 | 441 | 450 |

TABLE 3.—$N_0$ for r = 5

| N\b | 0 | 0.1 | 0.2 | 0.3 | 0.4 |
|---|---|---|---|---|---|
| 100 | 29 | 30 | 31 | 34 | 41 |
| 300 | 87 | 88 | 93 | 101 | 121 |
| 500 | 145 | 147 | 154 | 169 | 202 |
| 1000 | 290 | 294 | 308 | 337 | 403 |

TABLE 4.—$N_0$ for r = 10

| N\b | 0 | 0.1 | 0.2 | 0.3 | 0.4 |
|---|---|---|---|---|---|
| 100 | 23 | 24 | 25 | 28 | 34 |
| 300 | 70 | 71 | 74 | 82 | 100 |
| 500 | 116 | 118 | 124 | 137 | 167 |
| 1000 | 232 | 235 | 248 | 274 | 334 |

## REFERENCES


1. GEHAN E, and FREIREICH E. Non-randomized control in clinical cancer trials. N Engl J Med 290:198–203, 1974.
2. MEIER P. Statistics and medical experimentation. Biometrics 31:511–529, 1975.
3. POCOCK SJ. The combination of randomized and historical controls in clinical trials. J Chronic Dis 29:175–188, 1976.
4. POCOCK SJ. Randomized versus historical controls—a compromise solution. *In* Proceedings of the 9th International Biometric Conference. Raleigh, NC, Biometric Society, vol 1, 1976, pp 245–260.
5. KAPLAN EL, and MEIER P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481, 1958.
6. SUSARLA V, and VAN RYZIN J. Nonparametric Bayesian estimation of survival curves from incomplete observations. J Am Stat Assoc 71:897–902, 1976.
7. FERGUSON TS. A Bayesian analysis of some nonparametric problems. Ann Stat 1:209–230, 1973.
8. PETERSON AV. Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. J Am Stat Assoc 72:854–858, 1977.
9. BLUM J, and SUSARLA V. On the posterior distribution of a Dirichlet process given randomly right censored observations. Stocastic Processes and Their Applications 5:207–211, 1977.

# Comparisons of Cooperative Group Evaluation Criteria for Multiple-Drug Therapy for Breast Cancer[1]

**Hugh L. Davis, Jr,* Phyllis Multhauf, and Jerome Klotz[2]**

## SUMMARY

**The variability of response rates and durations for multiple-drug therapy for advanced breast cancer may be due in part to the evaluation criteria of the investigators. To assess the impact on outcome, a single data set comprising the patients registered in Central Oncology Group protocol 7020B was analyzed by the evaluation criteria of the Central Oncology Group, the Eastern Cooperative Oncology Group, the Southwest Oncology Group, and the Southeastern Cancer Study Group, as well as by a personally derived "flexible clinical interpretation of treatment benefit." The response rates and durations were quite similar by the criteria of all groups except those of the Southeastern Cancer Study Group, which gave lower rates and longer duration. These differences were consistent with the evaluation criteria for each group and confirmed the differences noted in their published reports. When "flexible" criteria were used, the differences deviated from those using strict group criteria. This might explain, in part, reports from single institutions of response rates that were higher than those achieved in cooperative group studies.**

[Cancer Treat Rep 64:507–517, 1980]

Fifteen years ago Greenspan et al (1) demonstrated increased response rates in advanced breast cancer using combinations of an alkylating agent, one or more antimetabolites, testosterone propionate, and prednisone. Greenspan (2) subsequently confirmed this observation in a larger series. In 1969 Cooper (3) combined 5-fluorouracil, methotrexate, cyclophosphamide, vincristine, and prednisone in what is now known as five-drug therapy or CMFVP, with a response rate of 90% in 88 patients. The plethora of clinical trials that followed dissected and refined these observations and utilized varying schedules of three to five of these agents. Recent reviews of these many trials by Carter (4), Tormey and Neifeld (5), and Carbone et al (6) stress that the average response rates are 50%, while noting a range of 20%–62% in individual series (5).

While agreeing in principle that some of these schedules could be more active than others, the duration of response and survival is quite similar, suggesting the influence of patient selection and evaluation criteria. To analyze the impact of evaluation criteria on response rates, criteria were chosen from published studies conducted between 1970 and 1974 by four major cooperative groups: Eastern Cooperative Oncology Group (ECOG), Southwest Oncology Group (SWOG), Southeastern Cancer Study Group (SEG), and Central Oncology Group (COG). Table 1 indicates the group, protocol number, drug doses and schedule, and the minimum adequate trial specified in the protocol. Table 2 documents published results of these studies in terms of response rates, duration, and survival. The single-agent comparisons in ECOG 0971, SWOG 450, and SEG 339, though important, will not be analyzed. Note the slightly greater response rates observed in the studies utilizing prednisone and continuous five-drug administration (COG 7020B, SWOG 450, and SEG 339) when compared to intermittent combination schedules (SWOG 450 and SEG 339) and a combination omitting prednisone (COG 7020B: CMFV). The low overall response rate of SEG 339 has puzzled oncologists and has been attributed to varying criteria of response and patient selection.

[1]Supported by Public Health Service grant CA-12271 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]University of Wisconsin, Madison.

**Reprint requests to*: H. L. Davis, Jr, MD, Division of Clinical Oncology, K4/648 Wisconsin Clinical Cancer Center, 600 Highland Ave, Madison, WI 53792.

TABLE 1.—Cooperative Group Studies involving combination chemotherapy in breast cancer

| Study | Schema* | | Minimum adequate trial |
|---|---|---|---|
| COG 7020B | 5-FU 10 mg/kg/wk iv (maximum dose: 500 mg/wk)<br>MTX 0.5 mg/kg/wk iv (maximum dose: 25 mg/wk)<br>VCR 0.02 mg/kg/wk iv (maximum dose: 1 mg/wk)<br>CTX 2 mg/kg/d orally (maximum 100 mg/d)<br>PRED 45 mg/d × 2 wk orally; *then* 30 mg/d × 2 wk; *then* 15 mg/d thereafter | *vs* 5-FU, MTX, VCR, CTX as in Regimen 1 | 8 wks |
| ECOG 0971 | 5-FU 600 mg/m²/iv Days 1, 8<br>MTX 40 mg/m²/iv Days 1, 8<br>CTX 100 mg/m²/orally daily × 14<br>Cycles repeated at 28-day intervals | *vs* Melphalan 6 mg/m²/d × 5 orally q6wks | 24 wks† |
| SWOG 450 | 5-FU 180 mg/m²/d × 5 iv<br>MTX 4 mg/m²/d × 5 iv<br>VCR 0.625 mg/m²/Days 1, 5 iv<br>CTX 120 mg/m²/d × 5 iv<br>PRED 40 mg/m²/d × 5 orally<br>Repeat courses q28days | *vs* 5-FU 300 mg/m²/wk iv<br>MTX 15 mg/m²/wk iv<br>VCR 0.625 mg/m²/wk iv<br>CTX 60 mg/m²/d orally<br>PRED 30 mg/m²/d × 14 orally; *then* 20 mg/m²/d × 14; *then* 10 mg/m²/d | 10 wks† |
| | | *vs* Adriamycin 60 mg/m²/iv q3wks (maximum total dose: 500 mg/m²) | 9 wks† |
| SEG 339 | 5-FU 400 mg/m²/wk iv (maximum dose: 1 g)<br>MTX 20 mg/m²/wk orally (maximum dose: 25 mg)<br>VCR 1 mg/m²/wk iv (maximum dose: 1 mg)<br>CTX 100 mg/d orally<br>PRED 45 mg/d (taper) | *vs* 5-FU 400 mg/m²/iv Days 1, 8<br>MTX 30 mg/m²/orally Days 1, 8<br>VCR 1 mg/m²/iv Days 1, 8<br>CTX 400 mg/m²/iv Day 1<br>PRED 20 mg 4 times/d × 7 orally<br>Cycles repeated at 28-day intervals | 8 wks† |
| | Agents administered sequentially until relapse on each | *vs* 5-FU 600 mg/m²/wk iv<br>MTX 20 mg/m²/twice weekly orally<br>VCR 1 mg/m²/wk iv<br>CTX 100 mg/m²/d orally<br>PRED 45 mg/d orally (taper) | 8 wks† (each agent) |

*5-FU = 5-fluorouracil; MTX = methotrexate; VCR = vincristine; CTX = cyclophosphamide; PRED = prednisone.
†Unless evidence of progression on therapy ensues; patients progressing after 4 wks of treatment were considered evaluable.

TABLE 2.—Analysis of published results of studies in breast cancer: multiple-drug therapy by 4 cooperative groups

| Study | Total No. of evaluable patients | No. of CRs (%) | No. of evaluable patients (%) with— | | Median duration (wks) of response (range) | Survival* rate of— | | | Ref No. |
|---|---|---|---|---|---|---|---|---|---|
| | | | PR | CR + PR | | 75% | 50% | 25% | |
| COG 7020B | | | | | | | | | |
| CMFVP | 48 | — | 30 (62.5) | 30 (62.5) | 18 (4–68+) | 31 | 54 | 93 | 7 |
| CMFV | 52 | — | 23 (44.2) | 23 (44.2) | | 31 | 60 | 106 | |
| ECOG 0971 | | | | | | | | | |
| CMF | 93 | 14 (15) | 35 (38) | 49 (53) | 24 (8–72+) | 25 | 52 | — | 8 |
| SWOG 450 (phase I) | | | | | | | | | |
| Weekly CMFVP | 106 | 19 (18) | 40 (38) | 59 (56) | 42 (4–112+) | — | ≈60 | — | 9 |
| Intermittent CMFVP | 98 | 8 (8) | 32 (32) | 40 (40.8) | 34 (4–104+) | — | ≈60 | — | |
| SEG 339 | | | | | | | | | |
| Continuous CMFVP | 35 | 4 (11) | 12 (35) | 16 (46) | 28 (4–78+) | ≈25 | 48 | ≈90 | 10 |
| Intermittent CMFVP | 33 | 4 (12) | 5 (15) | 9 (27) | 34 (15–115+) | ≈15 | 48 | ≈75 | |

*Survival values are in wks.

TABLE 3.—Comparative eligibility requirements

| Eligibility requirements | COG | ECOG | SWOG | SEG |
|---|---|---|---|---|
| Diagnosis | Histologic proof of breast carcinoma | Histologic proof of breast carcinoma | Breast carcinoma | Breast carcinoma |
| Age | At least 15 yrs | | | |
| Life expectancy | ≥ 3 mos | | | |
| Laboratory | Wbcs ≥ 4500/mm³<br>Platelets ≥ 150,000/mm³<br>BUN < 25 mg/100 ml | Wbcs ≥ 5000/mm³<br>Platelets ≥ 100,000/mm³<br>BUN < 25 mg/100 ml<br>Creatinine < 1.5 mg/100 ml | Normal renal function | |
| Measurable disease | Bidimensional tumor; counted lesions; biochemical abnormality (previously irradiated lesions not accepted for measurement) | Bidimensional lesion; excludes a) pleural effusion, b) mixed bone, c) irradiated lesions | Bidimensional lesion outside the field of prior radiation | Bidimensional lesion(s) |
| Prior therapy | | | | |
| Chemotherapy | Wait 4 wks with no continuing toxicity; must show definite evidence of progression prior to entry | None accepted | None accepted | None of the study agents |
| Hormones or steroids | No prior corticosteroids | Evidence of progression and 4-wk wait | Evidence of progression; started ≥ 2 mos before entry | 4-wk wait |
| Endocrine manipulation | Adrenalectomy or hypophysectomy excludes patient | 8-wk wait for surgical hormonal manipulation (12 wks for measurable osseous involvement) | 8-wk wait for irradiation to ovaries with evidence of progression | Premenopausal patients not acceptable unless previous ovarian ablation by surgery or radiation |
| Radiotherapy | Wait 4 wks; > 3000 rads to pelvis or lumbar spine excludes patient | Blood cell counts must be acceptable | Allowed if active disease outside the field | |
| Concomitant therapy | No concomitant drug therapy, hormones or palliative procedures allowed | | | |
| Other factors for exclusion | Any associated severe disease | | Myocardial disease randomized to combinations only | Premenopausal patients not acceptable unless previous ovarian ablation by surgery or radiation |

A comparison of patient selection introduces many intangibles, and meaningful interpretations are difficult. Comparison of criteria of eligibility, response, and evaluation is possible since strict criteria have been established for each protocol and reported in subsequent publications (7–10).

This report presents the results of a re-analysis of the patient population studied in COG 7020B. The patients were separately analyzed by the criteria established by each of the four groups. The factors that might have contributed to a variation in results were identified and the impact of these evaluation criteria assessed. This procedure may be pertinent to the analysis of current trials which have yet to be published.

## MATERIALS AND METHODS

The records of all patients registered in COG 7020B were independently reviewed by the strict criteria of the four groups as documented in COG 7020B, ECOG 0971, SWOG 450, and SEG 339, modified where necessary by the criteria given in the published results. The comparative eligibility criteria are shown in table 3 and the criteria of response are shown in table 4. In addition, a

TABLE 4.—Comparative evaluation criteria

| Measured lesion response* | COG† | ECOG | SWOG | SEG‡ |
|---|---|---|---|---|
| CR (complete response) | Complete disappearance of all objective evidence of tumor | Complete disappearance of all lesions and recalcification of skeletal lesions | Disappearance of all clinical evidence of active tumor; reossification of lytic lesions; absence of pain | No evidence of disease except for persistence of bone defects |
| PR (partial response) | ⩾ 50% decrease in the sum of the products of the diameters | ⩾ 50% decrease in the product of the largest diameters and/or partial reossification of lytic lesions | ⩾ 50% decrease in the sum of the products of the diameters; improvement in the skeleton; hypercalcemia controlled without aid | Good = ⩾ 50% decrease in *both* diameters of *all* measured lesions *or* ⩾ 75% reduction in the sum of the products<br>Fair = ⩾ 25% decrease in *both* diameters of *all* measured lesions *or* ⩾ 50% reduction in the sum of the products. Lytic lesions asymptomatic; do not increase in size |
| IMP (improvement) | | | Some reduction in tumor size but < 50%; disappearance or steady state of skeletal lesions on roentgenogram | |
| NC (no change) | Increase or decrease < 50% in sum of products of diameters | Decrease of < 50% or increase < 25% in sum of products of diameters; stable osseous lesions | Failure to improve with no obvious progression | Improvement in symptoms; lesions stable or with growth arrested after initiating therapy |
| Prog (progression)§ | ⩾ 50% increase in sum of products of diameters; new lesion | ⩾ 25% increase over original measurements in product of diameters; new lesions; progression of osteolytic lesions | ⩾ 50% increase of *any* measured lesion; new lesion; uncontrolled hypercalcemia and/or clearly progressive skeletal involvement | ⩾ 25% increase in the product of diameters of *any* lesion; new lesion |
| Progression after a response | Increase in the sum of the products over the minimal size which represents ⩾ 25% return to original size; new lesion | ⩾ 50% increase in product of diameters over minimal size; new lesion | ⩾ 50% increase in product of diameters of *any* measured lesion from minimal size; new lesion; reappearance of lesion after a CR; hypercalcemia; new or progressive skeletal involvement; abnormal clinical or lab observations secondary to disease | *Any* existing tumor progresses or new lesion appears |
| Response duration | | | | |
| PR | Time of 50% decrease to relapse | Time of 50% decrease to relapse | Time of 50% decrease to relapse | Time of response to relapse |
| CR | Time of NED¶ to relapse | Includes time in PR | Includes time in PR | Includes time in PR |

*A minimum duration of 4 wks required for any event in all groups.

†The COG was the only group to use counted lesions as a criterion for evaluation: CR = complete disappearance of all lesions being counted; PR = > 10% decrease in the No. of lesions being counted; NC = < 10% increase or decrease in the No. of lesions being counted; and Prog = > 10% increase in the No. of lesions being counted (any new lesion).

‡The SEG criteria were modified to take into account total tumor burden: CR = all traces of tumor disappeared; PR = 50% of the individual organs involved achieved a measurable response—lytic lesions become or remain asymptomatic and do not increase in size; NC = failure to achieve a PR; and Prog = appearance of a new lesion or any existing tumor achieved a progressive status.

§An initial evaluation of progression was assigned only when initial progression occurred within the first 4 wks of therapy.

¶NED = no evidence of disease.

TABLE 5.—Comparative response rates

| Denominator* | Total No. of patients included | Patients with— CR + PR No. | % | NC No. | % | Prog No. | % | Group |
|---|---|---|---|---|---|---|---|---|
| 1 | 106 | 46 | 43 | 31 | 29 | 6 | 6 | COG (A) |
| 2 | 102 | | 45 | | 30 | | 6 | |
| 3 | 83 | | 55 | | 37 | | 7 | |
| 4 | 88 | | 52 | | 35 | | 7 | |
| 1 | 109 | 43 | 39 | 35 | 32 | 7 | 6 | ECOG (B) |
| 2 | 105 | | 41 | | 33 | | 7 | |
| 3 | 85 | | 51 | | 41 | | 8 | |
| 4 | 91 | | 47 | | 38 | | 8 | |
| 1 | 116 | 58 | 50 | 38 | 33 | 3 | 3 | "Optimistic" (C) |
| 2 | 112 | | 52 | | 34 | | 3 | |
| 3 | 99 | | 59 | | 38 | | 3 | |
| 4 | 102 | | 57 | | 37 | | 3 | |
| 1 | 110 | 42 | 38 | 35 | 32 | 9 | 8 | SWOG (D) |
| 2 | 106 | | 40 | | 33 | | 8 | |
| 3 | 86 | | 49 | | 41 | | 10 | |
| 4 | 92 | | 46 | | 38 | | 10 | |
| 1 | 111 | 35 | 32 | 42 | 38 | 10 | 9 | SEG (E) |
| 2 | 107 | | 33 | | 39 | | 9 | |
| 3 | 87 | | 40 | | 48 | | 11 | |
| 4 | 93 | | 38 | | 45 | | 11 | |

*Denominators are defined in Results section.

category C was established as an optimistic flexible evaluation of overall treatment benefit as defined by H. L. Davis, Jr. The main focus will be on the strict criteria of the four groups; however, the "denominators" have been defined both in terms of the evaluable cases defined by the group criteria and in terms of three denominators suggested by a recent National Cancer Institute working group (11).

The percentage of response, no change, and progression, as well as the duration of response and time to progression, were calculated for the pooled patients of COG 7020B according to each group definition. The curves were calculated by the life-table method of Kaplan and Meier (12).

Comparisons of response criteria were carried out using a modification by Klotz (13) of the Cochran (14)-Friedman (15) statistic that permits missing observations and multiple-ordered categories of response.

## RESULTS

A tabulation of response (table 5) includes the following four different denominators:

1. total entered minus ineligible patients;
2. total entered minus ineligible and cancelled patients;
3. total entered minus ineligible, cancelled, unevaluable, early withdrawals due to drug toxicity, and early deaths; and
4. total evaluable patients as defined by each group's criteria (table 4).

Focusing on denominator 4, the varying response percentages could be easily explained:

(*a*) The COG had an additional category of "counted lesions" where a decrease of 10% in the number of lesions constituted a response. Two patients were in this category, which was not included in the other

TABLE 6.—Ordered categories of response

| Code | Categories |
|---|---|
| 0 | Ineligible, invalid, inadequate, incomplete, (missing category) |
| 1 | Complete response |
| 2 | Partial response, fair partial response, or good partial response |
| 3 | No change, stable, no response, no progression, or improvement |
| 4 | Progression or increasing disease |
| 5 | Early death |
| 6 | Drug toxicity |
| 7 | Drug death |

group criteria. Eliminating these two responses brings the COG response rates very close to those of the ECOG and SWOG.

(*b*) The low response rate of the COG patients analyzed by SEG criteria is noteworthy. The SEG demanded a response in at least 50% of the metastatic sites as an overall > 50% decrease of "tumor burden" (exclusive of osseous involvement). No other group employed this strict interpretation because there was little interest in the response of nonmeasured metastatic sites. Only stability of these additional sites was required. This restriction reduces the number and percent responding accordingly.

(*c*) Minor variations in criteria for fully evaluable patients accounted for denominator changes which were largest in the SEG and which further contributed to the lower response rates by this group's definition.

Category C, the most optimistic evaluation, was derived by flexibly interpreting eligibility criteria, such as prior chemotherapy, unacceptable laboratory values upon entry (eg, elevated BUN and low wbc count), counting late responses in the face of early progression, and ignoring minor progression in the face of overall clinical benefit. This flexibility boosts the response rates in the four denominators and may illustrate the effect of "loose" criteria on response rates.

A statistical analysis was carried out to compare the evaluation criteria using a modification (13) of Cochran's test (14), which permits use of more than two categories of response and handles missing observations in a particular category. Such missing observations arise because some patients are unacceptable under one criterion and not under another.

Appendix I defines the statistic, which is a quadratic form similar to the Friedman statistic (15), with modifications to handle missing observations (indicated by 0s). For the special case of no missing observations, it reduces to the Friedman statistic corrected for ties using average ranks (16). For only two categories of response and no missing observations, it is the Cochran statistic (14).

Table 6 shows the ordered categories of response used in the computation. Comparing all five criteria of evaluation (COG, ECOG, "optimistic," SWOG, and SEG), the value of the statistic is 15.46, which gives a $P$ value approximately 0.004 from a $\chi^2_4$ approximation. Eliminating the "optimistic" criteria, the categories are still significantly different (statistic 16.09 and $P$ value 0.001 from a $\chi^2_3$ approximation).

Figure 1 shows Kaplan-Meier estimates of the duration of response of *responders* by each of the definitions (A = COG, B = ECOG, C = "optimistic," D = SWOG, E = SEG). Note the wide separation between group C (approximately 12 weeks at the median) and the criteria of the four cooperative groups which show little variation except at the tail of the curves.

Figure 2 similarly shows estimates of time to progression from Day 1 of treatment for all evaluable patients. Again the curves are almost identical for all four groups, with only group C (optimistic) showing separation. A similar estimate of time to progression from Day 1 of treatment was made for responders only (not shown). The median duration of response was 21 weeks for COG, ECOG, and SWOG, compared to 18 weeks in figure 2. However, the SEG criteria show a 32-week median duration of response for this group (fig 2) consistent with the results of their own study (table 2), where the responses lasted longer because of more restrictive criteria of response (table 4).

## DISCUSSION

This analysis illustrates the variability of response rates obtained when a single data set is analyzed by different criteria. The near-identical duration of response and time-to-progression curves obtained suggests the similarity of these programs. This is further supported by similar survival figures (table 2). This analysis of response rates is consistent with the response rates reported by the groups for their own studies (table 2). The lower response rates reported by the SEG are consistent with their published results (10), and are strikingly similar to our data set (table 5). A more recent publication of this group gives an identical response rate (17), further documenting the consistency of their evaluation criteria.

The highest response rates and durations were ob-

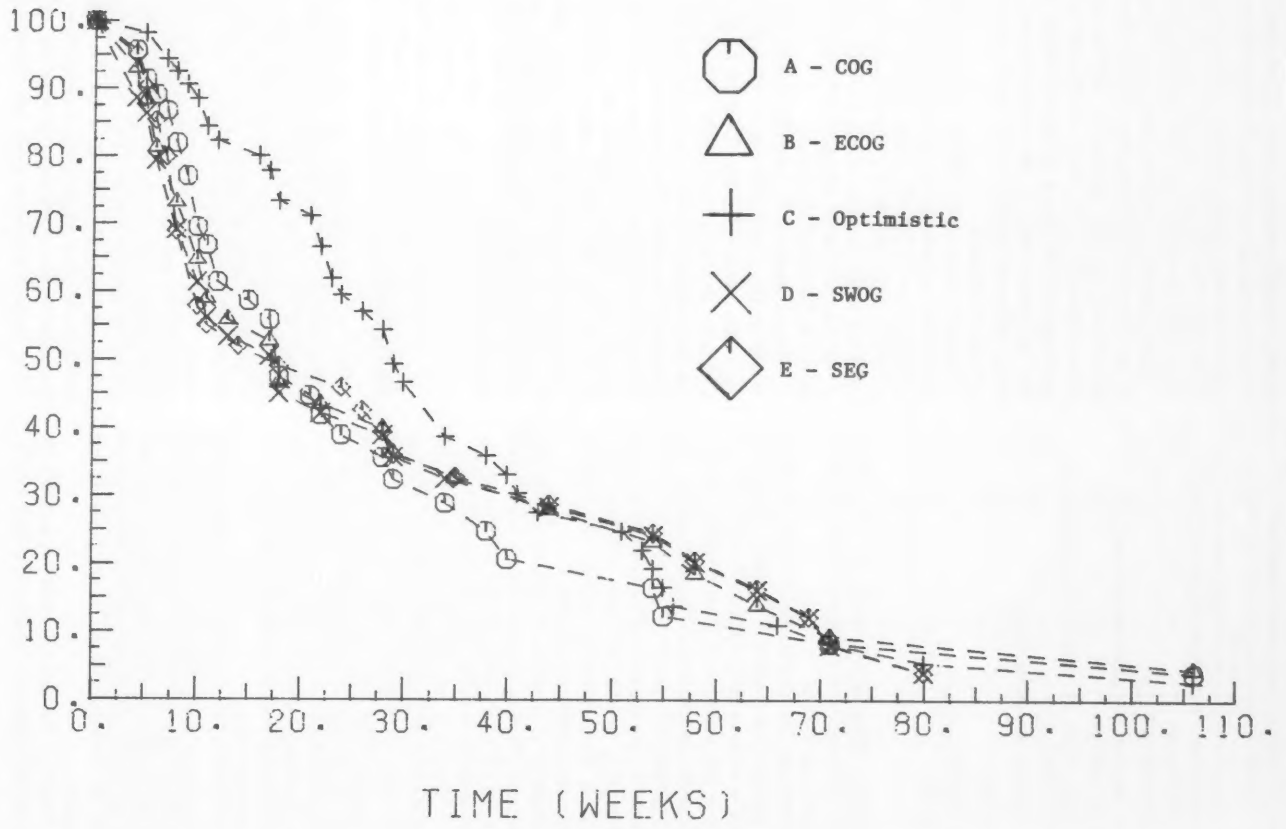



FIGURE 1.—Kaplan-Meier estimates of the duration of response of responders by each cooperative group and the flexible "optimistic" definition.

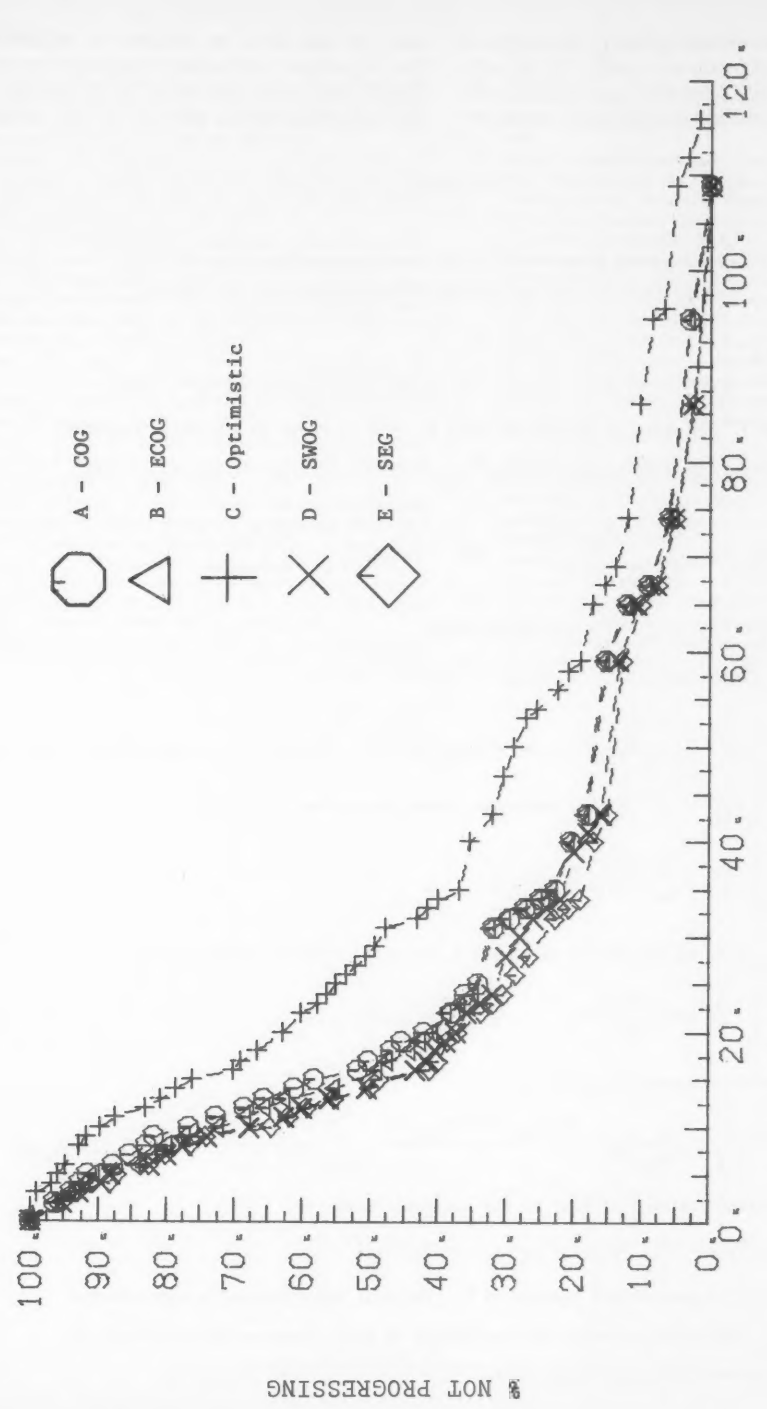



FIGURE 2.—Estimates of time to progression from Day 1 of treatment for all evaluable patients.

tained by using flexible criteria (group C) targeted to a subjective evaluation of treatment benefit. We can only speculate how often these somewhat loose criteria have been applied to single institution studies and so may account for high rates not obtained by the groups (4,5). Such investigator enthusiasm is dampened by application of strict criteria as employed in the group studies, which reflects a more realistic appraisal of these programs.

## APPENDIX

### Definition of a Modified Cochran-Friedman Statistic for Missing Observations and Multiple-Ordered Response Categories

Notation

Let $X_{ij}$ $(i = 1,2,\ldots,I,\ j = 1,2,\ldots,J)$ be the response of the $i^{th}$ patient for the $i^{th}$ category of evaluation with $X_{ij} = 0$ if there is a missing response. For distinct nonzero response values $z_{j1} < z_{j2} < \ldots < z_{jK_j}$ of $X_{1j}, X_{2j}, \ldots, X_{Ij}$ let

$$U_{ijk} = \begin{cases} 1 & \text{if } X_{ij} = z_{jk} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad U_{ijo} = \begin{cases} 1 & \text{if } X_{ij} = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Define

$$R_{ij} = \begin{cases} 0 & \text{if } U_{ijo} = 1 \text{ (missing response)} \\ U_{ij1}(U_{+j1}+1)/2 + U_{ij2}(U_{+j1} + (U_{+j2}+1)/2) + \\ \qquad \ldots + U_{ijk_j}(U_{+j1}+U_{+j2} + \ldots + (U_{+jK_j}+1)/2) & \text{otherwise (average rank)} \end{cases}$$

$$R_{i+} = \sum_{j=1}^{J} R_{ij} \qquad \text{(+ indicates summation on the subscript)}$$

$$E_{i+} = \sum_{j=1}^{J} (I - U_{+jo} + 1)(1 - U_{ijo})/2$$

$\ddagger = (\sigma_{ii'})$ is an IxI matrix of rank I-1 where the diagonal elements are

$$\sigma_{ii} = \sum_{j=1}^{J} \left\{ \frac{(I-U_{+jo}+1)(I-U_{+jo}-1)}{12} - \sum_{k=1}^{K_j} \frac{(U^3_{+jk}-U_{+jk})}{12(I-U_{+jo})} \right\} (1-U_{ijo})$$

and the off diagonal elements are

$$\sigma_{ii'} = -\sum_{j=1}^{I} \frac{1}{(I-U_{+jo}-1)} \left\{ \frac{(I-U_{+jo}+1)(I-U_{+jo}-1)}{12} - \sum_{k=1}^{K_j} \frac{(U_{+jk}-U_{+jk})}{12(I-U_{+jo})} \right\} (1-U_{ijo})(1-U_{i'jo}).$$

The statistic is then defined by the quadratic form

$$K = (R_{1+}-E_{1+}, \ldots, R_{I+}-E_{I+})\ \ddagger^{-}\ (R_{1+}-E_{1+}, \ldots, R_{I+}-E_{I+})'$$

where $\ddagger^{-}$ is a generalized inverse of $\ddagger$. The null distribution is approximated by a $\chi^2_{I-1}$ distribution under the hypothesis of equal response probabilities for the different evaluation categories.

## REFERENCES


1. Greenspan EM, Fieber M, Lesnick G, et al. Response of advanced breast carcinoma to the combination of the antimetabolite methotrexate and the alkylating agent Thio-Tepa. J Mt Sinai Hosp NY 30:246–267, 1963.
2. Greenspan EM. A combination and sequential chemotherapy approach to advanced mammary carcinoma. *In* Progress in Clinical Cancer (Ariel IM, ed). New York, NY, Grune and Stratton, 1973, vol 5, pp 81–105.
3. Cooper RG. Combination chemotherapy in hormone-resistant breast cancer. Proc Am Assoc Cancer Res 10:15, 1969.
4. Carter SK. Chemotherapy of breast cancer: current status. *In* Breast Cancer: Trends in Research and Treatment (Heuson JC, Mattheiem WH, and Rozencweig, eds). New York, NY, Raven Press, 1976, pp 193–215.
5. Tormey DC, and Neifeld JP. Chemotherapeutic approaches to disseminated disease. *In* Breast Cancer Management—Early and Late (Stoll BA, ed). London, England, William Heinemann, 1977, pp 117–131.
6. Carbone PP, Bauer M, Band P, et al. Chemotherapy of disseminated breast cancer—current status and prospects. Cancer 39:2916–2922, 1978.
7. Ramirez G, Klotz J, Strawitz JG, et al. Combination chemotherapy in breast cancer. A randomized study of 4 versus 5 drugs. Oncology 32:101–108, 1975.
8. Canellos GP, Pocock SJ, Taylor SG III, et al. Combination chemotherapy for metastatic breast carcinoma—prospective comparison of multiple drug therapy with L-phenylalanine mustard. Cancer 38:1882–1886, 1976.
9. Hoogstraten B, George SL, Samal B, et al. Combination chemotherapy and Adriamycin in patients with advanced breast cancer—a Southwest Oncology Group Study. Cancer 38:13–20, 1976.
10. Smalley RV, Murphy S, Huguley CM, Jr, et al. Combination versus sequential five-drug chemotherapy in metastatic carcinoma of the breast. Cancer Res 36:3911–3916, 1976.
11. Hoogstraten B (Chariman). Breast cancer proposed guidelines. Report from the Combination Chemotherapy Trials Working Group. *In* Breast Cancer Task Force Treatment Committee, National Cancer Institute. Breast Cancer: Suggested Protocol Guidelines for Combination Chemotherapy Trials and for Combined Modality Trials. Washington, DC, US Department of Health, Education, and Welfare (DHEW). Public Health Service National Institutes of Health (NIH)—DHEW Publication (NIH) 77–1192.
12. Kaplan EL, and Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481, 1958.
13. Klotz JH. A Modified Cochran-Friedman Test With Missing Entries and Ordered Categorical Data. Technical Report No. 522. Madison, WI, Department of Statistics, University of Wisconsin, 1978.
14. Cochran WG. The comparison of percentages in matched samples. Biometrika 37:256–266, 1950.
15. Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J Am Stat Assoc 32:675–701, 1937.
16. Lehmann EL. Nonparametrics: Statistical Methods Based on Ranks. San Francisco, CA, Holden Day, 1975, p 265.
17. Smalley RV, Carpenter J, Bartolucci A, et al. A comparison of cyclophosphamide, Adriamycin, 5-fluorouracil (CAF) and cyclophosphamide, methotrexate, 5-fluorouracil, vincristine, prednisone (CMFVP) in patients with metastatic breast cancer. Cancer 40:625–632, 1977.

# Design of Phase II Clinical Trials in Cancer Using Decision Theory[1]

**Richard J. Sylvester[2,*] and Maurice J. Staquet[3]**

## SUMMARY

**An application of decision theory to the design of phase II clinical trials in cancer is presented. A risk function which depends on the prior distribution of the response rate, the costs of treating a patient, and the gains or losses resulting from the decisions taken upon completion of the study is derived and optimal restricted sampling plans for a response rate of 20% are provided.**

[Cancer Treat Rep 64:519–524, 1980]

Decision theory is not widely used in medicine because of the difficulties encountered in the formulation of a realistic model that is easy to work with and acceptable to statisticians and clinicians alike. Expository articles (1,2) have advocated the use of decision theory as an aid to patient management and treatment, and other papers have described its use in coronary artery surgery (3) and radiotherapy (4), and in the treatment of cancer (5,6) and renal disorders (7–9). Dunnett (10) has proposed a decision theoretic approach to the screening of new drugs by pharmaceutical companies. Decision theory has also been used to determine which of two or more treatments is best (11–18). This method has not been used in practice, partly because it provides for a selection between treatments based on a loss function rather than on a test of a hypothesis (19).

In phase II trials, however, selection procedures rather than tests of hypotheses are appropriate. In cancer, phase II clinical trials screen new drugs in humans to detect possible antitumor activity. In theory, each trial treats ten to 30 patients with measurable evolutive disease in a range of signal tumor types. For each tumor type studied it is important to decide with as few patients as possible whether:

*a.* a new drug is possibly active in a specified proportion of patients and thus additional studies should be undertaken to assess further the effectiveness of the drug; or

*b.* a new drug is probably inactive and should be dropped from further study in a given tumor type.

In practice, the number of patients treated in a phase II trial may vary considerably from one trial to another and is often determined by the number of patients with a given type of tumor who become available for treatment during a given period. Gehan (20) has, however, proposed a sampling plan based on the laws of probability. The most serious error in screening trials is to reject an effective drug from further study; Gehan's plan controls this error (type II error or false negative) by calculating the probability that n consecutive patients do not respond to the given treatment for a prespecified rate of response to the drug. The sample size is determined as the smallest value of n for which the probability of n consecutive failures is less than some given error rate. If 0 of n patients respond to the drug, the drug is rejected from further study. If at least one of n patients responds to the drug, additional patients are treated in order to attain a given standard error of the response rate.

Gehan's plan is simple to use, but it has the following drawbacks.

*a.* At the end of a phase II study, a wide confidence interval for the response rate of drugs which are active is invariably left.

*b.* A priori information concerning the possible effectiveness of the new drug or the response rates of other new drugs in previous phase II studies of a given tumor type is not taken into account.

*c.* The (ethical) costs of treating a patient with a new drug or the gains or losses resulting from the decisions taken at the completion of the phase II study are not considered.

[1]Supported by Public Health Service grant CA-11488 from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]European Organization for Research on Treatment of Cancer Data Center, Institut Jules Bordet, Brussels, Belgium.

[3]European Organization for Research on Treatment of Cancer Data Center, Institut Jules Bordet, and Laboratoire de Statistique Médicale, Faculté de Médecine, Université Libre de Bruxelles, Belgium.

**Reprint requests to:* Dr. Richard J. Sylvester, European Organization for Research on Treatment of Cancer Data Center, Institut Jules Bordet, Brussels, Belgium.

The results of our preliminary investigation into the use of decision theory in phase II trials (21,22) provide an introduction to the concepts and terminology used here.

The purpose of this paper is to describe a decision theory approach to the design of phase II clinical trials in which the information provided by the prior distribution of the response rate, the costs of treating a patient, and the gains or losses resulting from the decisions taken are used. A risk function which determines the optimal number of patients to be treated and the minimum number of responses required before the drug can be accepted for further study is presented.

## THE MODEL

For each type of tumor studied let:

- $n$ = the number of patients treated with the new drug in the phase II study.
- $x$ = the number of patients in the phase II study who respond to the new drug.
- $c$ = the rejection number. If $x \leq c$, the drug is rejected from further study while if $x > c$, the drug is accepted for further study.
- $P$ = the response rate of the new drug.
- $P_0$ = the lowest rate of response that is considered to be of practical importance. If $P > P_0$, then the treatment of patients with the new drug would constitute a gain, while if $P < P_0$, the treatment of patients with the new drug would constitute a loss. For $P = P_0$, the loss in treating patients with the new drug is considered to be equal to zero. $P_0$ may vary according to the tumor site, whether or not the patients have been previously treated with other drugs, and what other treatments (if any) are available.
- $m$ = the number of patients treated with the new drug in the subsequent study if the drug is found to be active in the phase II study. Although we shall assume that the next study is a comparative phase III trial, the results which follow are valid, whether or not this is true.
- $t$ = the average number of patients who are treated with an effective new drug after completion of the phase II study before a second new drug which is at least as good is found.
- $k$ = the difference in (ethical) cost of further treatment (with a possibly different treatment regimen) between a patient who does not respond to the new drug and a patient who does respond. We will assume that k is positive.

Using the above definitions, the loss function, as shown in table 1, can be derived. There is a loss, $k(P_0-P)n$, due to treating n patients with a new drug in a phase II study, and another loss depending on the decision taken at the conclusion of the study. The overall loss is then the sum of these two losses. Note that if $P > P_0$ and the drug is accepted for further study, the overall loss is negative and a gain has been realized.

If we let $R(P)$ = the probability that we reject the new drug from further study

$$= \sum_{x=0}^{c} \binom{n}{x} p^x (1-p)^{n-x}$$

and

$A(P)$ = the probability that we accept the drug for further study

$$= \sum_{x=c+1}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$

then one can write the risk, the expected value of the overall loss, as

$$R(n, c, P) = \begin{cases} k(P_0-P)n + k(P_0-P)mA(P) & P < P_0 \\ 0 & P = P_0 \\ k(P_0-P)n + 2k(P-P_0)tR(P) + k(P_0-P)t & P > P_0 \end{cases}$$

The optimal design is given by the values of n and c, which minimize the risk $R(n,c,P)$. Since the optimal values of n and c depend on the unknown value of P, no one plan is optimal for all values of P. One method of eliminating the dependence on P is to specify a prior distribution $w(P)$ for P, which represents the investigator's prior beliefs concerning the true therapeutic effectiveness of the drug in a given type of tumor. For the purpose of this paper we shall assume that P can take on two values $P_1$ and $P_2$ with probabilities $w_1$ and $w_2$, respec-

TABLE 1.—Loss function for phase II clinical trial

| True therapeutic effectiveness | Action | | |
|---|---|---|---|
| | During phase II study— | At conclusion of phase II study— | |
| | Treat n patients | Accept drug for further study | Reject drug from further study |
| $P < P_0$ | $k(P_0-P)n$ | $k(P_0-P)m$ | 0 |
| $P = P_0$ | 0 | 0 | 0 |
| $P > P_0$ | $k(P_0-P)n$ | $k(P_0-P)t$ | $k(P-P_0)t$ |

tively, $w_1 + w_2 = 1$ and $P_1 < P_0 < P_2$. Thus, $P_1$ represents the response rate of a drug that should be rejected from further study, and one would want $A(P_1)$ to be small, whereas $P_2$ represents the response rate of a drug that should be accepted for further study, and one would want $R(P_2)$ to be small. In practice $P_2$ should be chosen to be the lowest rate of response greater than $P_0$ for which one wishes to have a high probability of accepting the drug for further study. The part of the risk function depending on n and c can then be written as:

$$R(n,c) = (P_0 - E(P))n + \gamma_1 A(P_1) + \gamma_2 R(P_2)$$

where $E(P) = P_1 w_1 + P_2 w_2$ is the expected rate of response before the start of the phase II study,

$$\gamma_1 = w_1(P_0 - P_1)m \quad \text{and}$$
$$\gamma_2 = 2w_2(P_2 - P_0)t$$

The optimal values of n and c minimizing the risk R(n,c) are independent of the cost constant k, but do depend on $P_1$, $P_2$, $w_1$, $w_2$, m, and t. In practice, one may also be interested in determining the optimal restricted sampling plan where a restriction of the form $R(P_2) = \beta$ is placed on the probability of rejecting an effective drug with response rate $P_2$ from further study.

## RESULTS

This paper presents only a summary of some of the most important results; the details will be reported elsewhere.[4]

If $E(P) > P_0$, the risk decreases as n increases, so that no optimal plan exists. In this case the drug is expected to have a higher response rate than the minimal response rate of interest, and the sample size should be as large as possible. For $E(P) < P_0$, the case of most practical importance, a minimum will exist. Since most new drugs tested in phase II trials cannot be expected a priori to have a response rate greater than the minimum response rate of interest, in what follows we shall assume that $E(P) < P_0$, so that an optimal plan will exist.

For unrestricted sampling plans where no restriction is placed on $R(P_2)$, there is no exact explicit solution for both n and c in terms of $P_1$, $P_2$, $w_1$, m, and t. However, for large samples it can be shown that:

(*a*) c is a linear function of n, so that c increases as n increases;
(*b*) n is proportional to log m, so that as the size of the phase III study which follows increases, the size of the phase II trial should also increase;
(*c*) n is approximately proportional to log t; and
(*d*) n decreases as $w_1$ increases.

If, however, one makes the restriction $R(P_2) = \beta$, then the sample size n is independent of t. (In order for the relation $R(P_2) = \beta$ to be satisfied exactly, it is necessary to randomize between two successive values for n.) Examples of optimal restricted sampling plans are given in tables 2 and 3 for the special case where $P_0 = 0.20$, $P_2 = 0.25$, and $R(P_2) = 0.05$.

## Early Stopping Rules

Ethical considerations could prompt one to close a study based on a single sampling plan before the required number of patients have been treated. Suppose, for example, that the optimal sampling plan is to take n = 34 and c = 4. If, in fact, the drug is not active and no

TABLE 2.—Restricted single sampling plans for $P_0 = 0.20$, $P_2 = 0.25$, and $R(P_2) = 0.05$

| Possible values (for n ≤ 74) of —* | | |
|---|---|---|
| n | v | c |
| 10 | 0.5515 | 0 |
| 11 | 0.4485 | |
| 17 | 0.9894 | 1 |
| 18 | 0.0106 | |
| 22 | 0.0696 | 2 |
| 23 | 0.9304 | |
| 28 | 0.4667 | 3 |
| 29 | 0.5333 | |
| 33 | 0.0953 | 4 |
| 34 | 0.9047 | |
| 39 | 0.8468 | 5 |
| 40 | 0.1532 | |
| 44 | 0.7001 | 6 |
| 45 | 0.2999 | |
| 49 | 0.6398 | 7 |
| 50 | 0.3602 | |
| 54 | 0.6477 | 8 |
| 55 | 0.3525 | |
| 59 | 0.7121 | 9 |
| 60 | 0.2879 | |
| 64 | 0.8255 | 10 |
| 65 | 0.1745 | |
| 69 | 0.9829 | 11 |
| 70 | 0.0171 | |
| 73 | 0.1617 | 12 |
| 74 | 0.8383 | |

*Where v is the probability of choosing the corresponding value of n for a given c.

[4]Sylvester RJ. A Bayesian approach to the design of phase II clinical trials in cancer. Manuscript in preparation.

patients respond, a minimum of 30 patients would have to be treated before the study could be stopped and the drug could be declared ineffective. Some type of sequential approach is desirable so that ineffective drugs could be rejected from further study before the planned completion of the study. Since patients may be entered in groups and the response to treatment is usually not known before the next patient's entry in the study, a fully sequential approach is not possible. However, it is possible to stop the study after a certain number of patients have been entered in order to take a preliminary look at the results. Depending on the results obtained, one of the three following decisions could be made.

1. Stop the phase II study and reject the drug from further study.
2. Continue the phase II study to its normal completion.
3. Stop the phase II study and begin a phase III study immediately.

With a doubling sampling plan based on the above principles, ineffective drugs will be quickly rejected from further study. The number of patients required by the double sampling plan can be restricted so as not to exceed the number of patients required by the single sampling plan. This will eliminate the possibility that the double sampling plan will lead to undesirably large sample sizes for drugs which are moderately effective, and will ensure that effective drugs are sent to phase III trials as soon as possible, and before results become so promising that a phase III trial becomes unethical to carry out.

TABLE 3.—Optimal restricted single sampling plans for $P_0 = 0.20$, $P_2 = 0.25$, and $R(P_2) = 0.05$*

| | m = 50 | | m = 75 | | m = 100 | | m = 125 | | |
|---|---|---|---|---|---|---|---|---|---|
| $P_1$ | c | $A(P_1)$ | c | $A(P_1)$ | c | $A(P_1)$ | c | $A(P_1)$ | $w_1$ |
| | † | † | † | † | † | † | † | † | 0.01 |
| | † | † | † | † | † | † | † | † | 0.10 |
| | 1 | 0.0123 | 1 | 0.0123 | 1 | 0.0123 | 2 | 0.0015 | 0.25 |
| 0.01 | 0 | 0.0997 | 1 | 0.0123 | 1 | 0.0123 | 1 | 0.0123 | 0.50 |
| | 0 | 0.0997 | 1 | 0.0123 | 1 | 0.0123 | 1 | 0.0123 | 0.75 |
| | 0 | 0.0997 | 1 | 0.0123 | 1 | 0.0123 | 1 | 0.0123 | 0.90 |
| | 0 | 0.0997 | 1 | 0.0123 | 1 | 0.0123 | 1 | 0.0123 | 0.99 |
| | † | † | † | † | † | † | † | † | 0.01 |
| | † | † | † | † | † | † | † | † | 0.10 |
| | † | † | † | † | † | † | † | † | 0.25 |
| 0.05 | 2 | 0.1045 | 3 | 0.0521 | 3 | 0.0521 | 3 | 0.0521 | 0.50 |
| | 1 | 0.2080 | 2 | 0.1045 | 3 | 0.0521 | 3 | 0.0521 | 0.75 |
| | 1 | 0.2080 | 2 | 0.1045 | 2 | 0.1045 | 3 | 0.0521 | 0.90 |
| | 1 | 0.2080 | 2 | 0.1045 | 2 | 0.1045 | 3 | 0.0521 | 0.99 |
| | † | † | † | † | † | † | † | † | 0.01 |
| | † | † | † | † | † | † | † | † | 0.10 |
| | † | † | † | † | † | † | † | † | 0.25 |
| 0.10 | 5 | 0.1927 | 7 | 0.1156 | 8 | 0.0891 | 9 | 0.0685 | 0.50 |
| | 2 | 0.4061 | 4 | 0.2478 | 6 | 0.1495 | 7 | 0.1156 | 0.75 |
| | 2 | 0.4061 | 4 | 0.2478 | 5 | 0.1927 | 6 | 0.1495 | 0.90 |
| | 1 | 0.5185 | 3 | 0.3180 | 5 | 0.1927 | 6 | 0.1495 | 0.99 |
| | † | † | † | † | † | † | † | † | 0.01 |
| | † | † | † | † | † | † | † | † | 0.10 |
| | † | † | † | † | † | † | † | † | 0.25 |
| 0.15 | † | † | † | † | † | † | † | † | 0.50 |
| | 0 | 0.8164 | 4 | 0.5897 | 9 | 0.3970 | 12 | 0.3108 | 0.75 |
| | 0 | 0.8164 | 0 | 0.8164 | 4 | 0.5897 | 8 | 0.4303 | 0.90 |
| | 0 | 0.8164 | 0 | 0.8164 | 1 | 0.7478 | 6 | 0.5043 | 0.99 |
| | † | † | † | † | † | † | † | † | 0.01 |
| | † | † | † | † | † | † | † | † | 0.10 |
| | † | † | † | † | † | † | † | † | 0.25 |
| 0.19 | † | † | † | † | † | † | † | † | 0.50 |
| | † | † | † | † | † | † | † | † | 0.75 |
| | 0 | 0.8888 | 0 | 0.8888 | 0 | 0.8888 | 4 | 0.7987 | 0.90 |
| | 0 | 0.8888 | 0 | 0.8888 | 0 | 0.8888 | 0 | 0.8888 | 0.99 |

*For each value of c given in this table, the corresponding values of n and v are given in table 2.

†Since $E(P) \geq P_0$, the risk decreases as the values of n and c given in table 2 increase. Hence, the optimal sampling plan in practice is to choose the largest feasible values of n and c as given in table 2.

## AN EXAMPLE

Suppose that for a given type of tumor the lowest response rate of interest is 20%, and that if the response rate is as high as 25%, one wants to restrict the probability of rejecting the drug from further study to 0.05. Before the study starts, the investigator has reason to believe that there is a 50% chance that the response rate may be as low as 5%, and a 50% chance that the response rate may be as high as 25%. Then, if the number of patients who would eventually be treated with the new drug in a phase III study is 75, we find from table 3 (with $P_1 = 0.05$, $w_1 = 0.50$, and $m = 75$) that the optimal plan is to take $c = 3$, and from table 2 that one should take $n = 28$ with probability 0.4667 and $n = 29$ with probability 0.5333. If in practice we do not wish to randomize, we can take $n = 29$ as an approximation which will be accurate enough for all practical purposes. If, on the other hand, the investigator believes that there is at least a 75% chance that the response rate may be only 5%, then the optimal value of n is approximately 23 and $c = 2$.

These plans may be curtailed through the use of double sampling plans in order to avoid treating a large number of patients with an ineffective drug. Examples of such plans will be given in a forthcoming paper (*see* footnote 4).

## DISCUSSION

Since the purpose of phase II trials is to select or reject drugs for further study based on only a small number of patients, decision theoretic procedures are believed to be more appropriate than the classic procedures of hypothesis testing and estimation which are used in phase III trials and which require large numbers of patients.

In the design of a phase II study based on decision theory, the following procedure should be followed for each type of tumor studied.

1. Select $P_0$, the minimum response rate of interest.
2. Select a prior distribution for the response rate. That is, select two response rates $P_1$ and $P_2$, $P_1 < P_0 < P_2$, and their respective probabilities $w_1$ and $w_2$. In practice, $P_2$ should be chosen to be close to $P_0$ and may in fact be chosen before choosing $P_0$ when a restricted sampling plan is used.
3. Select m, the number of patients who eventually would be treated with the new drug in a phase III study.
4. Decide whether to restrict the probability of rejecting an effective new drug with response rate $P_2$ from further study. If no restriction is made, specify the value of t, the average number of patients treated with an effective new drug after completion of the phase II study before a second new drug, which is at least as good is found.
5. Determine the values of n and c which minimize the risk function.
6. Determine an early stopping rule so that the study can be stopped before its planned completion if the drug turns out to be completely inactive.

As in any trial in which no control group is chosen by randomization, it is most important to report as accurately as possible the patient entry criteria defined in the protocol and the characteristics of those patients actually entered in the study. Only in this manner can the results of such a study have any meaning in evaluating the drug's effectiveness.

## REFERENCES


1. Kassirer JP. The principles of clinical decision making: an introduction to decision analysis. Yale J Biol Med 49:149–164, 1976.
2. Pauker SG, and Kassirer JP. Therapeutic decision making: a cost-benefit analysis. N Engl J Med 293:229–234, 1975.
3. Pauker SG. Coronary artery surgery: the use of decision analysis. Ann Intern Med 85:8–18, 1976.
4. Lusted LB. Decision-making studies in patient management. N Engl J Med 284:416–424, 1971.
5. Henschke UK, and Flehinger BJ. Decision theory in cancer therapy. Cancer 29:1819–1826, 1967.
6. Safran C, Desforges JF, and Psichlis PN, eds. Diagnostic Planning and Cancer Management. Massachusetts Institute of Technology, Laboratory for Computer Science, Clinical Decision Making Group, Cambridge, and the Blood Research Laboratory, Tufts New England Medical Center Hospital, Department of Medicine, Tufts University School of Medicine, Boston, MA, 1976.
7. Gorry GA, Kassirer JP, Essig A, et al. Decision analysis as the basis for computer-aided management of acute renal failure. Am J Med 55:473–484, 1973.
8. McNeil BJ, Varady PD, Burrows BA, et al. Measures of clinical efficacy: cost-effectiveness calculations in the diagnosis and treatment of hypertensive renovascular disease. N Engl J Med 293:216–221, 1975.
9. Schwartz WB, Gorry GA, Kassirer JP, et al. Decision analysis and clinical judgement. Am J Med 55:459–472, 1973.
10. Dunnett CW. Statistical theory of drug screening. *In* Quantitative Methods in Pharmacology (de Jonge H, ed). Amsterdam, The Netherlands, North-Holland Publishing Co, 1961, pp 212–231.
11. Anscombe FJ. Sequential medical trials. JASA 58:365–383, 1963.
12. Canner PL. Selecting one of two treatments when the responses are dichotomous. JASA 65:293–306, 1970.
13. Colton T. A model for selecting one of two medical treatments. JASA 58:388–400, 1963.
14. ———. A two-stage model for selecting one of two treatments. Biometrics 21:169–180, 1965.
15. Cornfield J, Halperin M, and Greenhouse SW. An adaptive procedure for sequential clinical trials. JASA 64:759–770, 1969.
16. Day NE. Two stage designs for clinical trials. Biometrics 25:111–118, 1969.
17. Donner A. The use of auxiliary information in the design of a clinical trial. Biometrics 33:305–314, 1977.
18. Mendoza G, and Iglewicz B. A three phase sequential model for controlled clinical trials. Biometrika 64:201–206, 1977.

19. SIMON R. Adaptive treatment assignment methods and clinical trials. Biometrics 33:743–749, 1977.

20. GEHAN EA. The determination of the number of patients required in a preliminary and follow up trial of a new chemotherapeutic agent. J Chronic Dis 13:346–353, 1961.

21. STAQUET M, and SYLVESTER R. A decision theory approach to phase II clinical trials. Biomedicine 26:262–266, 1977.

22. SYLVESTER RJ, and STAQUET MJ. An application of decision theory to phase II clinical trials in cancer. *In* Recent Advances in Cancer Treatment (Tagnon HJ, and Staquet MJ, eds). New York, NY, Raven Press, 1977, pp 1–11.

# Pilot Studies and New Agent Programs of the Eastern Cooperative Oncology Group[1]

John M. Bennett[2,*] and Marvin Zelen[3]

## SUMMARY

**The Eastern Cooperative Oncology Group (ECOG) has established two mechanisms for the study of new agents, including group A and B drugs and innovations in radiation therapy, namely, "pilot studies" and "master protocols." The master protocols allow for rapid accumulation of cases of tumor types of low incidence and the speedy completion of a study of new agents in common tumors. A new agents and pilot studies committee has been organized with responsibility for educational sessions, the introduction of new agents into either pilot studies or master protocols, and the review and monitoring of these programs. During its two years of operation, > 51 new studies have been initiated.**

[Cancer Treat Rep 64:525–532, 1980]

The Eastern Cooperative Oncology Group (ECOG) consists of 28 full and probationary member institutions with a current annual accrual of approximately 3700 cases. Most of these accessions involve randomized prospective phase III trials, 20% of which are of the combined modality type. Between 35 and 40 protocols are usually active at any one time, with about 25% considered to be phase II studies.

In recent years the study of phase II agents (so-called Group B drugs) has been carried out utilizing the "master protocol" schema (fig 1). Each protocol consists of two parts, a master section (Part A) and drug treatments (Part B) (fig 2). This allows for a convenient, rapid substitution of a brief addendum for each new specific agent when the current studies are nearing completion. A protocol schema is illustrated in figure 3.

In general, patients who are eligible for these protocols will have received conventional or combination therapy (phase III studies) and have measurable disease. Where, however, no effective chemotherapy exists or the objective response rate is < 15% in a given tumor type such as renal cell carcinoma, the introduction of Group B agents in previously untreated patients with generalized disease appears reasonable and rational.

Currently there are six master protocols involving the following tumor types: lymphomas, non-oat-cell lung cancer, renal cell carcinoma, multiple myeloma, melanoma, and ovarian cancer. In general, the practice in these studies has been to obtain 20–40 adequately treated patients in each tumor category. A group of 20 evaluable patients showing no response indicates that if activity for a given agent actually exists, the response rate is, at most, 14% (one-sided 95% confidence limit) (1). In general, one can quickly obtain an upper limit of the uncertainty of an observer response by adding $0.8/\sqrt{n}$ to the observed proportion of the responses. For example, if one observes four regressions among 40 treatments, the theoretic upper bound is $0.10 + 0.8/\sqrt{40}$ or 0.23 or a maximum response rate of 23% with an approximate 95% confidence limit.

It is important that the characteristics both of patients and of their disease be considered in such studies. The ECOG experience has demonstrated that, for advanced disease, the single most important variable that affects response is the pretreatment performance status. Patients who are nonambulatory or bedridden for ≥ 50% each day (ECOG 3–4) have little or no chance of a meaningful response with active therapies (2).

[1]Supported by Public Health Service grants CA-11083, CA-11198 (University of Rochester), and CA-12721, from the National Cancer Institute, National Institutes of Health, Department of Health, Education, and Welfare.

[2]Division of Clinical Services, University of Rochester Cancer Center and the Department of Medicine, University of Rochester School of Medicine and Dentistry, NY.

[3]Sidney Farber Cancer Institute, Division of Biostatistics and Epidemiology, Harvard University School of Public Health, Boston, MA.

**Reprint requests to:* John M. Bennett, MD, University of Rochester Cancer Center, 601 Elmwood Ave, Rochester, NY 14642.

EASTERN COOPERATIVE ONCOLOGY GROUP

PHASE II

MASTER PROTOCOL FOR EVALUATION OF NEW AGENTS IN PATIENTS WITH LYMPHOMA

FIGURE 1.—Front page of ECOG master protocol.

Obviously, other variables such as previous treatment, age, stage of disease, and histology can affect response. Theoretically, ignoring these variables may result in the accrual of patient populations that obscure activity. On the other hand, it is difficult to justify (on practical grounds) committing large numbers of patients to phase II treatments that may not have more than modest activity.

In an attempt to balance these conflicting concerns an experimental plan for a series of phase II trials of new agents was initiated in 1974. The advantages of studying new agents by this mechanism in a cooperative group include first, the ability to amass a considerable number of cases of tumor types of relatively low incidence (table 1). In 1974 a randomized phase II study in advanced renal cell cancer (EST 1874) involving the four agents Megace, cyclophosphamide, VP-16-213, and galactitol was started. One hundred and sixty-one patients were randomized, and the patient accession rate was approximately seven patients per month during a 2-year active accrual. The study was closed when 30 patients had entered each arm, and an abstract of the results was presented at the American Society of Clinical Oncology meeting in 1977. B. Lemley, Registry Coordinator for the University of Rochester Cancer Center's Regional Tumor Registry, has calculated that it would have taken > 6 years to achieve the same number of accessions for a population base at risk of approximately 700,000/year.

The second advantage is the completion of a study of new agents in common tumors with large accruals in a short period of time. An example of this type of activity can be found in EST 2575, The Master Protocol for Evaluation of New Treatments in Patients with Bronchogenic Carcinoma. This program was launched in October 1975, examining both single new agents and combinations of drugs on a minimum of 20 patients in each cell type (fig 4). If a response was observed in the first 20 patients, an additional five patients were allocated to the particular therapy up to a maximum of 40 patients (ie, four or more responses). Over the next 22 months 505 patients were entered in the study, with sufficient accrual to close most arms, including the crossover treatment with ascorbic acid and VM-26, by December 1977.

TABLE 1.—Data on master protocol EST 1874: phase II studies for renal cell carcinoma

| | |
|---|---|
| Activated | Dec 1974 |
| Terminated | Dec 1976 |
| Agents studied | Megace, cyclophosphamide, VP-16-213, galactitol |
| No. of patients— | |
| Randomized | 161 |
| Ineligible | 5 |
| Evaluable as of Dec 1976 | 126 |

A disadvantage of multi-institutional participation through the master protocol mechanism is that overestimation of the potential accrual may prolong the study several years beyond what was predicted before activation. This occurred in EST 5472, A Study of Adriamycin and Bleomycin in Patients with Multiple Myeloma. Only 32 patients were eventually registered, although originally it had been estimated that more than twice as many would be available for the study. Another problem relates to the use of the "crossover schema," which allows the same patient to be treated later with the agent or agents that he or she did not receive at initial randomization. The possibilities of cross-resistance and changes

PART A MASTER PROTOCOL

- 1.0 Background
- 2.0 Objectives
- 3.0 Patient Selection
- 4.0 Mechanics of Study
- 5.0 Induction Treatment
- 6.0 Definition of Terms
- 7.0 Remission Maintenance
- 8.0 Radiation or Corticosteroid Therapy of Patient While on Study
- 9.0 Measurement of Effect
- 10.0 Study Parameters
- 11.0 Pathology Review
- 12.0 Patient Consent and Peer Judgment
- 13.0 Records to be Kept

PART B. DRUG TREATMENTS

- 14.0 Schema and Details of Drug Treatment and Randomization
  - 14.1 Bleomycin
  - 14.2 VM-26

FIGURE 2.—Index for master protocol (EST 1474).

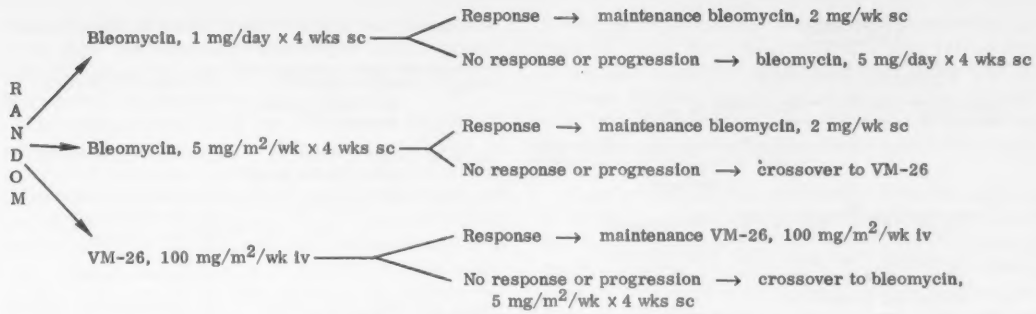

RANDOM
Bleomycin, 1 mg/day x 4 wks sc
Response → maintenance bleomycin, 2 mg/wk sc
No response or progression → bleomycin, 5 mg/day x 4 wks sc
Bleomycin, 5 mg/m²/wk x 4 wks sc
Response → maintenance bleomycin, 2 mg/wk sc
No response or progression → crossover to VM-26
VM-26, 100 mg/m²/wk iv
Response → maintenance VM-26, 100 mg/m²/wk iv
No response or progression → crossover to bleomycin, 5 mg/m²/wk x 4 wks sc

FIGURE 3.—Schema for master protocol (EST 1474).

in performance status should be considered as stratification variables, but are often ignored in such studies.

Although the ECOG had established a Therapeutic Modality Committee, it had done little to provide access to new agents or to monitor pilot studies. In 1977 I was asked by P. Carbone, *Chairman* of the ECOG, to reorganize this committee. The purpose was twofold: 1. to coincide with the decision by the Division of Cancer Treatment (DCT) of the National Cancer Institute to refine for the cancer centers and cooperative groups (table 2) the list of available groups for the study of phase II agents; and 2.—a complementary decision—to allow and encourage pilot studies of either phase II or III type under the aegis of a cooperative group.

The administrative structure of the committee consists of a chairman, an executive secretary, an advisory steering committee (with a three-member pharmacology subcommittee), and a committee of the whole consisting of representatives from both disease-oriented committees and member institutions (fig 5). By attending phase I-II drug liaison meetings sponsored by the DCT and by establishing contact with private pharmaceutical houses, a mechanism for rapid exchange of information was developed (fig 6).

Abstracts of pertinent material of Group A and B drugs are prepared quarterly and distributed via the ECOG Operations Office to all disease-oriented chairpersons and principal investigators. Representatives from the DCT

EST 2575 (Phase II Therapies)

EPIDERMOID

Institution Group 1
- E: Cyclophosphamide + CCNU
- J: Actinomycin D
- M: Methotrexate + Adriamycin + cyclophosphamide + CCNU

Institution Group 2
- K: Galactitol
- L: Mitomycin C
- N: Cyclophosphamide + *cis*-dichlorodiammineplatinum(II) + bleomycin

ADENOCARCINOMA

Institution Group 1
- M: Methotrexate + Adriamycin + cyclophosphamide + CCNU
- O: Dibromodulcitol + Adriamycin
- P: Ftorafur

Institution Group 2
- E: Cyclophosphamide + CCNU
- Q: Piperazinedione
- R: Adriamycin + 5-fluorouracil + *cis*-dichlorodiammineplatinum(II)

LARGE CELL

Institution Group 1
- E: Cyclophosphamide + CCNU
- Q: Piperazinedione
- R: Adriamycin + 5-fluorouracil + *cis*-dichlorodiammineplatinum(II)

Institution Group 2
- E: Cyclophosphamide + CCNU
- O: Dibromodulcitol + Adriamycin
- P: Ftorafur

DISEASE PROGRESSION:

RANDOM
- H: VM-26
- I: Ascorbic acid

FIGURE 4.—Schema for EST 2575: master protocol for bronchogenic carcinoma.

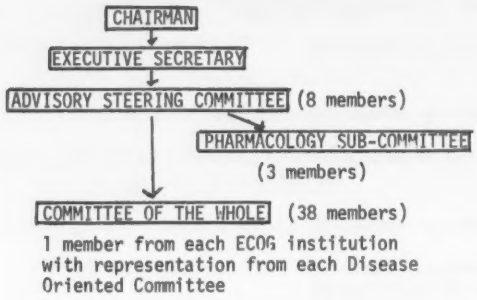



FIGURE 5.—ECOG NAPSC administrative structure.

FIGURE 6.—Exchange of information diagram for Group A and B agents.

TABLE 2.—Grouping of drugs qualifying for protocol study (modified from DCT handout)

| | | |
|---|---|---|
| Group A | Drugs in phase I and limited phase II studies | (1) Phase I/II working group |
| Group B | Drugs in extended phase II and phase III studies | (1) Phase I/II working group<br>(2) Phase III contractors<br>(3) Cooperative groups<br>(4) Study groups<br>(5) Task forces<br>(6) Cancer centers |
| Group C | Drugs with proven efficacy for specific indications | (1) Individual investigators |

are asked to make presentations at the New Agents and Pilot Studies Committee (NAPSC) educational sessions at the semi-annual ECOG national meetings. The stimulation for a pilot study can derive from the committee, the Disease Oriented Committee (DOC), or an individual investigator. New agents can be studied group-wide, as previously, through the master protocol or via the mechanism of having one to four institutions restrict their activities in one tumor type for advanced patients to a Group A or B drug. One example of the latter is EST P-A477, a pilot study for the use of maytansine in patients with lymphoma (fig 7). In contrast to previous phase II

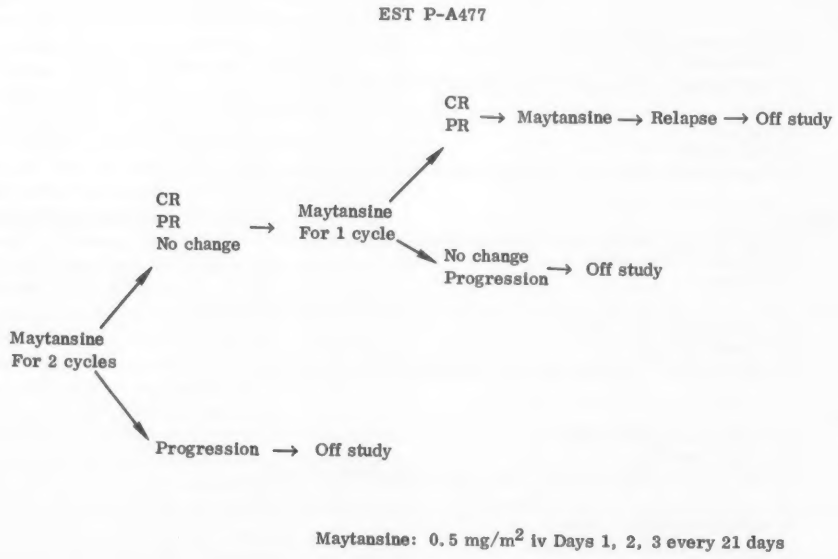



FIGURE 7.—EST P-A477: schema for maytansine lymphoma study.

studies in lymphomas, where accruals have been relatively slow, the enthusiasm shown by a limited number of investigators for this new approach has produced 14 accessions in < 5 months.

The major objectives of the NAPSC are to prevent unnecessary duplication of studies being conducted by others and to prevent excessive proliferation of studies by any one institution. Guidelines have been established for pilot studies, including the identification of toxicity and efficacy of single drugs or combinations of agents in one to four institutions with ≤ 30 patients, and the limitation of the number of active pilots to one to three per disease or tumor category. Figure 8 outlines the mechanism for implementation of a new protocol. The study is prepared using regular ECOG protocol guidelines and is submitted to the ECOG Operations Office for distribution to both the NAPSC and DOC. Either committee (or both) can reject the protocol outright, suggest revisions, or approve it. The chairman and executive officer assume much of the responsibility for rapid review and confer with the DOC by telephone. Members of the Steering Committee receive copies of the protocol and have 14 days to respond. If approved, the study receives a pilot study code number, is typed in final form, and is submitted to the Clinical Investigations Branch of the DCT for approval. Once approved, the study is distributed and activated. The average time from submission of the proposed new study to approval and activation has been 10 weeks. Since its inception in early 1977, a total of 100 pilot studies have been submitted to the NAPSC. Of these, 51 have been activated, 25 have been rejected, and the remainder are in various stages of processing. As of December 1979, > 700 patients have been entered in these protocols.

Since several of these pilot studies would be expected to produce positive results that could be incorporated into future ECOG protocols, a mechanism for review of these accessions was necessary. Because of the limited scope of each study and the small numbers involved, the Statistical Center with its cohort of data managers and statistical study coordinators is not primarily involved in these protocols. Each investigator files an onstudy form and interim reports and submits final forms to the ECOG Operations Office. Distribution is made to a designated NAPSC reviewer as well as to the study chairperson (fig 9). A case evaluation summary form (fig 10) is prepared 1 month before each ECOG meeting and is included in the minutes. Disagreements between the NAPSC reviewer and the pilot study chairman in regard to eligibility criteria, toxicity, and response are resolved by oral or written communication, and the final evaluation is sent to the Operations Office for filing and distribution to the DOC. Credit for patient accessions is then awarded to the participating institutions. Close monitoring of the

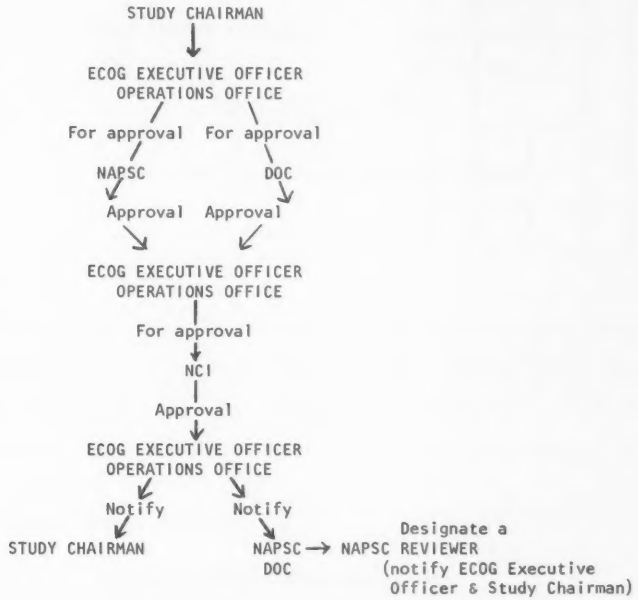



FIGURE 8.—Flow chart for implementation of ECOG pilot studies.

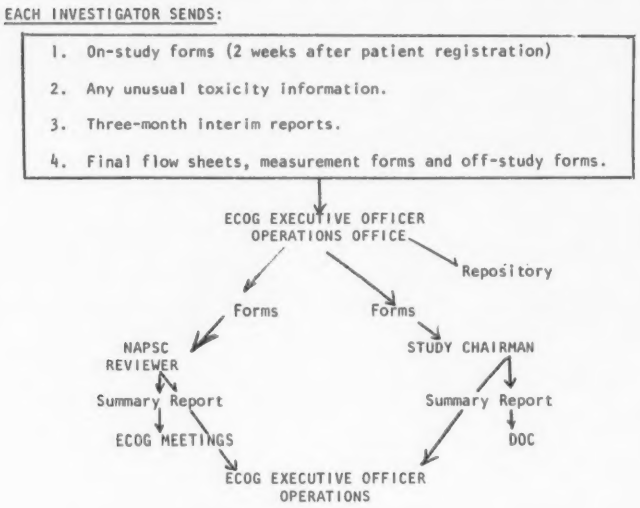



FIGURE 9.—Flow chart for investigator responsibilities.

ECOG NEW AGENTS AND PILOT STUDIES CASE EVALUATION SUMMARY FORM

The following study _____activated on _____ chaired by Dr.(s) _____has been reviewed by Dr. _____on _____.

During the past six (6) months _____patients have entered this study making a total patient entry of _____with the following results:

Number of patients evaluable for RESPONSE = _____

Number of RESPONSES (complete & partial) = _____

Number of patients evaluable for TOXICITY = _____

| | ECOG Toxicity Grade | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 |
| HEMATOLOGY | | | | | |
| G.I. | | | | | |
| NEUROLOGIC | | | | | |
| G.U. | | | | | |
| INFECTION | | | | | |
| OTHER (specify) | | | | | |
| | | | | | |
| | | | | | |

COMMENTS:

_____

_____

_____

_____

Case Evaluation Form (Pilot) 7/77

FIGURE 10.—Case evaluation summary form.

accessions is carried out, and if 75% of the required number has not accumulated within 1 year of activation, a termination of the protocol is recommended.

In summary, a new committee, organized within the ECOG, has been charged with the responsibility of stimulating the introduction of new agents and new pilot studies and monitoring these programs for toxicity and responses. During the first year of operation, close liaison with the disease-oriented chairpersons and staff members of the DCT has produced a steady flow of new pilot studies. A longer period of time will be necessary to determine how successful this mechanism can be in furthering the goal of identifying promising new leads in clinical cancer therapy.

## REFERENCES


1. GEHAN EA. Early human studies with anticancer agents: the question of sample size. J Chronic Dis 15:265–268, 1962.
2. ZELEN M. Importance of prognostic factor in planning therapeutic trials. *In* Cancer Therapy: Prognostic Factors and Criteria of Response (Staquet MJ, ed). New York, NY, Raven Press, 1975.

# Summary Views: A Statistician's Perspective

D. R. Cox [1,*]

## SUMMARY

**This paper outlines some broad issues concerning the role of statistical methods, and argues for flexibility of approach. It reviews the controversies concerning historical versus concurrent controls, and randomization versus extensive balancing. Finally, it proposes directions in which further developments of statistical technique would be desirable.**

[Cancer Treat Rep 64:533–535, 1980]

The organizers of the conference chose the title of these remarks; the word "perspective" is particularly apt in that it emphasizes dependence on the position of the observer. My perspective is not that of a medical statistician but of a theoretic statistician interested in applications, especially those in physical sciences and technology.

Time does not permit a detailed comment on individual papers. A number of the papers with no direct statistical content were especially impressive, and I apologize to their authors for appearing to ignore their contributions.

In discussing statistical issues it is sometimes useful to make a rough distinction between tactics and strategy. Tactics are concerned with details of design and data collection, with the choice of specific methods of analysis and their implementation on the computer, and, finally, with the presentation of conclusions in a way that will be both vivid and honest. Strategy is about broader issues of design and analysis, including, for example, the achievement of a reasonable balance of effort between data collection and data analysis, the role of historical data, and the establishment of a link with underlying scientific information. The following comments are mainly about strategic issues.

The role of specialists, such as statisticians, raises general issues in advanced societies. Nowadays there are few fields of science and technology in which the individual primarily responsible for an investigation handles all aspects of the work. In statistical matters, it seems highly desirable that the primary investigator and the statistician have genuine discussion, at least of the strategic issues. Involvement in more detailed aspects will depend on the individuals concerned; for example, those relatively few physicists who discuss their work with statisticians are likely to be interested in technical statistical details.

The constraints on investigation imposed by ethical considerations have inevitably figured quite prominently; for example, they are crucial in the assessment of Dr. Zelen's ingenious suggestion of randomized consent designs. To put the issue in perspective, it should be noted that constraints on investigation are common in many fields and that ethical problems also arise in other subjects, for instance in studies of recidivism in criminology. Almost universally, one central aspect of statistical design is the implementation of basic principles as intelligently as possible in the light of existing constraints.

All this is related to the need for flexibility in both design and analysis. For example, the great advantages of concurrent controls need to be emphasized continually; nevertheless, careful study of methods for the use of historical controls, and for the combined use of historical and concurrent controls, is to be welcomed. Recent work in statistics has put some welcome emphasis on flexible exploratory analysis. While there is usually need in the design stage for specification of the techniques of analysis to be adopted, this should not preclude intelligent modification and further analysis when the data become available. The following reasonable approach provides protection against bias in selecting a method of analysis: use one or more methods of analysis specified in advance, and always report the conclusions. If, in the light of the data, some other analysis seems preferable, present the case for this explicitly as part of the report.

Personal qualities of judgment, enthusiasm, and commitment play an important part in research work. With-

[1]Department of Mathematics, Imperial College, London, England.

**Reprint requests to:* Dr. D. R. Cox, Department of Mathematics, Imperial College of Science and Technology, Huxley Bldg, Queen's Gate, London SW7 2BZ, England.

out enthusiasm, work of value is unlikely to be started, let alone completed. The public character of science and the necessity for conclusions to carry broad conviction serve partly to control uncritical enthusiasm, and the use of standardized methods of statistical analysis plays a key role in this control. It is, however, important that statistical methods are not seen as primarily negative in character. Related to the theme of the personal element in research is the role of subjective Bayesian methods which have received so much attention in statistical circles in recent years and which were used in two of the investigations reported at the present meeting. Here the investigator's judgment, for example his a priori judgment of the reliability of historical controls, is entered as a quantitative value to achieve a synthesis of prior information and that provided by the data. It seems to me fruitful to study such ideas for qualitative insight, but I am deeply skeptical of the wisdom of using subjective probability quantitatively in the present context; this would be to treat on an equal footing information from the data and less clearly based information. To my mind, they are best kept separate. However, to take this view is not to undervalue the role of personal judgment in research work.

The important general problem of assembling conclusions from different sources has broad implications in statistics because there is a tendency, for obvious reasons of simplicity, to analyze data in isolated pieces. The (objective) Bayesian formalism is likely to be valuable here; several papers at the present conference touched, at least indirectly, on this topic.

The role of significance tests and the sensible use of $P$ values came up a number of times in discussion. A quite widely held view among statisticians now is that, while tests have their uses, they are often considerably overemphasized. Only rarely, for example, should a significance test by itself be the main focus of the analysis.

## Some Requirements for a Good Investigation

The one overwhelming requirement for a good investigation is that there be an important idea to study. Conceivably, this obvious fact may be overlooked if too much emphasis is placed on technique. The object of a study is normally formalized as the comparison of two or more "treatments," interpreting this word in a very broad sense. It is important that the treatments be precisely standardized; where, as in surgery, this may not be practicable, the main characteristics of the treatment should be recorded in each case.

The next requirement is that there be satisfactory measures of response. Sometimes survival will be the most important measure of response, although even then the investigator should bear in mind the possibility of recovering information from supplementary response variables. This is especially so when the study contains an appreciable number of "censored" individuals who may yield a measurable supplementary response. In other cases, where there is no definitive measure of response, the work on psychosocial measures seems especially welcome.

Finally, there is the choice of explanatory (or prognostic) variables, both to increase the accuracy of comparison and to detect interaction with treatment effects. This aspect is briefly discussed below.

## Two Key Issues

Two more technical issues concern historical versus concurrent controls and, given the use of concurrent controls, the use of randomization versus extensive balancing.

As noted earlier, other things being equal, the arguments for concurrent controls are overwhelming; however, when the number of available patients is limited or when the investigator has relatively large quantities of high-quality recent historical data, the best option may be critical use of historical data instead of or as well as concurrent controls. Having said this, I should emphasize the need for further empirical evidence on the stability of comparisons involving historical controls.

Data from historical controls can be "adjusted" for comparison with (current) treated individuals by relatively elementary methods or by the more elaborate procedures studied by Dr. Gehan. In either case, the main qualitative assumptions involved are:

1. that the treatment applied in the historical control period is, in all its aspects, the one that is appropriate for comparison with the new treatment, ie, that general technological advance has not made the historical control treatment irrelevant; and
2. that relevant explanatory variables ignored in the adjustment procedure maintain the same relation with the included variables. Ignored variables are of two kinds: those recorded but excluded from the adjustment procedure and those not measured. The validity of assumption 2 for the first kind of ignored variable can and should be checked.

The issue of randomization versus extensive balancing, ie, the use of systematic designs, has been under discussion in one form or another ever since explicit randomization was first suggested as a technique of experimental design. Two special features of the present applications are their sequential character and the particular need to avoid selection bias on entry, for which some form of randomization seems highly desirable. Drs. Simon and Brown have reviewed the main arguments carefully, and so I will make only the following brief comments.

*a.* Simplicity is very important.

*b.* The average loss of efficiency in adjusting for explanatory variables by analysis rather than by bal-

ancing depends on the ratio of the number of stratification parameters to the number of individuals.

c. Subject to (a), a sensible strategy may be to force near balance on explanatory variables thought very likely to have strong effects, but to prevent extreme imbalance on all potential explanatory variables.

Concerning more technical statistical matters, Dr. George discussed the role of sequential tests. In principle, sequential tests with delayed response do not seem to pose any difficulty, although there is the embarrassing possibility that the conclusion, when the data are finally complete, is different from that indicated when the entry of new individuals was stopped; in extreme cases this would be evidence of heterogeneity of treatment effect. Formal sequential stopping rules demand a single criterion as a basis for stopping; more detailed analysis at the end of an investigation will usually ignore the sequential character of the investigation. If it is possible that stopping is based on an apparent effect in a subset of individuals, it is very important that the study design incorporate appropriate allowances for selection; otherwise, premature stopping is virtually inevitable.

More technical research is needed on adjustment procedures. Should some adaptation of ridge regression be used? Are automatic variable selection procedures at all appropriate? Are allowances for selection needed when there are many potential explanatory variables? Will it be wise to include more explanatory variables than would be normal in an interpretative use of regression? How can assumptions be checked?

Finally, there is scope for further work on statistical design and analysis when treatments consist of complex sequences of component treatments, possibly modified in light of intermediate responses.

# Perspective of the Clinician

**Barth Hoogstraten** [1,*]

The topic of this symposium "Designs for Clinical Cancer Research" is very timely, and we have heard some excellent presentations. "Some" of course implies that a number of papers were of lower caliber and a few were of questionable relevance to this conference.

The tone of the symposium was set by the fact that the majority of presentations were made by statisticians. The use of historical controls remains a controversial issue for nearly all statisticians and clinicians. Dr. Gehan is its principal advocate, and he bears the brunt of the attack. It should be remembered that he does not favor wide use of the historical control, but rather a limited one. Its use should not be the debatable issue but rather the definition of "limited."

Dr. Carter stated that the "pilot idea" is central to the design of trials, but I always believe that the "question" is the most essential ingredient—especially the ability to ask a question. The idea to answer the question comes later.

The importance of prognostic factors in the design of clinical trials was stressed by several speakers, and in this area the Cooperative Groups have been of great importance. The careful design of on-study forms and disease-specific flow sheets has made the analysis of prognostic factors possible. The groups provide numbers of patients large enough to elicit statistically significant differences of factors which heretofore were considered to be of lesser importance and which now are being used in the design of studies for patient subgroups.

Dr. Miller hid a fine presentation behind an obscure title. He enumerated lymphoblast morphology, immunoglobulin levels, and marrow status on Day 14 as important factors in the design of trials. The Children's Cancer Study Group has led the way in finding these factors. Other examples of sophisticated studies come to mind. The differentiation of kappa and lambda chains early in the evaluation of myeloma studies by the Cancer and Acute Leukemia Group B, and the testing for B, T, null cell, and even pre-B cell leukemia by the Southwest Oncology Group are just two examples of what the Cooperative Groups can accomplish.

Statistical presentations can be mind-boggling to the clinician. Statisticians speak their own language, but for the sake of better clinical trials it is necessary that statistician and clinician understand each other, and I urge the clinical statistician to work in that direction. At the same time I hope that my clinical colleagues will hide less behind the *P* value, especially the premature and unconfirmed *P* value. And the biostatisticians in their analysis of ongoing trials would do well to incorporate time as a factor for designating the statistically significant importance of the statistically significant difference. Time has a way of changing the *P* value, and should be in the calculations.

The need for at least some standardization of surgical techniques was clearly demonstrated by Dr. DeCosse. It remains to be seen to what extent the normally fiercely independent surgeons are willing to accept this.

Evidently, this is a lesser problem for the radiotherapists. Already inclined to mathematics and strict dosimetry, the quality control of radiotherapy may be easier to establish. *There* the danger may be that quality control may lead to the tail wagging the dog.

Most impressive were the presentations by Drs. Leventhal and Weiss on immunotherapy. Dr. Leventhal emphasized the many problems facing purification of biologic agents and safety of whole cell and cell fraction vaccines; that transfer factor is still more theory than fact, and that the lessons learned from the animal model, ie, need for a small tumor and application of BCG close to the tumor, were not followed in human studies. Dr. Weiss delivered a most scientific and hard-hitting speech in which he strongly condemned the current practice of immunotherapy. Dosage is wrong and all too often too high; the time of application in relation to chemotherapy is based on a feeling of what may be correct rather than on good research, and the animal model cannot be used. Yet thousands of patients throughout the world are receiving BCG, *Corynebacterium parvum*, and levamisole from their physicians, who in good faith follow the example set by some investigators who should know better but who prefer both the limelight and the additions to their curriculum vitae. The message from both speakers was clear: stop and go back to the drawing board. It seems like good advice and should make one think twice before designing future combined modality studies with immunotherapy.

Psychosocial studies of cancer were initially conducted by amateur psychologists, ie, oncologists who were genuinely concerned about the psychosocial problems of can-

[1]Chairman, Southwest Oncology Group, Kansas City, KS.

**Reprint requests to*: Barth Hoogstraten, MD, Chairman, Southwest Oncology Group, Rosedale Bank, Rainbow Boulevard, Kansas City, KS 66103.

cer and who tried to fill the void left by the professionals. It is refreshing to observe that more psychiatrists are becoming interested in this area, even to the extent that they refer to themselves as psychosocial oncologists.

Drs. Vietti and Herson touched on a distinct weakness. The Cooperative Groups have a major impact on cancer therapy throughout the world, and it seems only a small step for the Groups to adopt uniform toxicity criteria. It is long overdue and the differences between the Groups are small, especially for chemotherapy. Reason should replace undue sense of independence. And while they are working at that, the Groups might as well tackle the toxicity criteria for other treatment modalities, especially those for radiotherapy.

The criteria of response used by the Cooperative Groups are now almost identical. What are not identical are the users of these criteria. And there lies the weakness in percentage of response. Criteria have never led to standardization among investigators, who vary from pessimistic to overoptimistic. For that reason I tend to disregard early response rates, give a jaundiced eye to final response rates, view with some suspicion remission duration curves, and put more stock in survival curves. No amount of calculation or evaluation or analysis has ever changed duration of life.

Finally, my evaluation of the statisticians at large: as a Group Chairman, I am exposed to surgeons, pathologists, pediatricians, radiotherapists, immunologists, and chemotherapists, and I have learned to appreciate that they did not wander into their specialty by chance. Their psychologic make-up made them drift into their field, almost unavoidably. Our exposure these 3 days to so many statisticians has taught me that they too form a special breed. It is highly statistically significant, $P$ value $< 0.0035$, that so many wear a beard. There is little heterogeneity among these; the variance is small, since these beards are short and neatly clipped. Their Bayesian survival is as yet unknown because there are too many censored survivors of these beards. But seriously, they are driven with great tenacity toward the truth. It was a revealing experience and it makes me admire even more the tremendous contributions made by the biostatistical oncologists to the treatment of cancer. If only they could learn to control the time limit of their talks.



☆ U.S. GOVERNMENT PRINTING OFFICE: 1980 O— 311-209/9

If your address has changed and you wish to continue receiving this publication, please send your new address and the mailing label from this copy to:

Cancer Treatment Reports
National Cancer Institute
Landow Bldg, Rm 2C16
7910 Woodmont Ave
Bethesda, MD 20014

(301) 496-6975

If you do not wish to continue receiving this publication, please notify us at the address noted above and enclose the mailing label from this copy.

A publication of the National Cancer Institute

NIH Publication No. 80-22